



The ECMWF ensembles

Roberto Buizza

European Centre for Medium-range Weather Forecasts

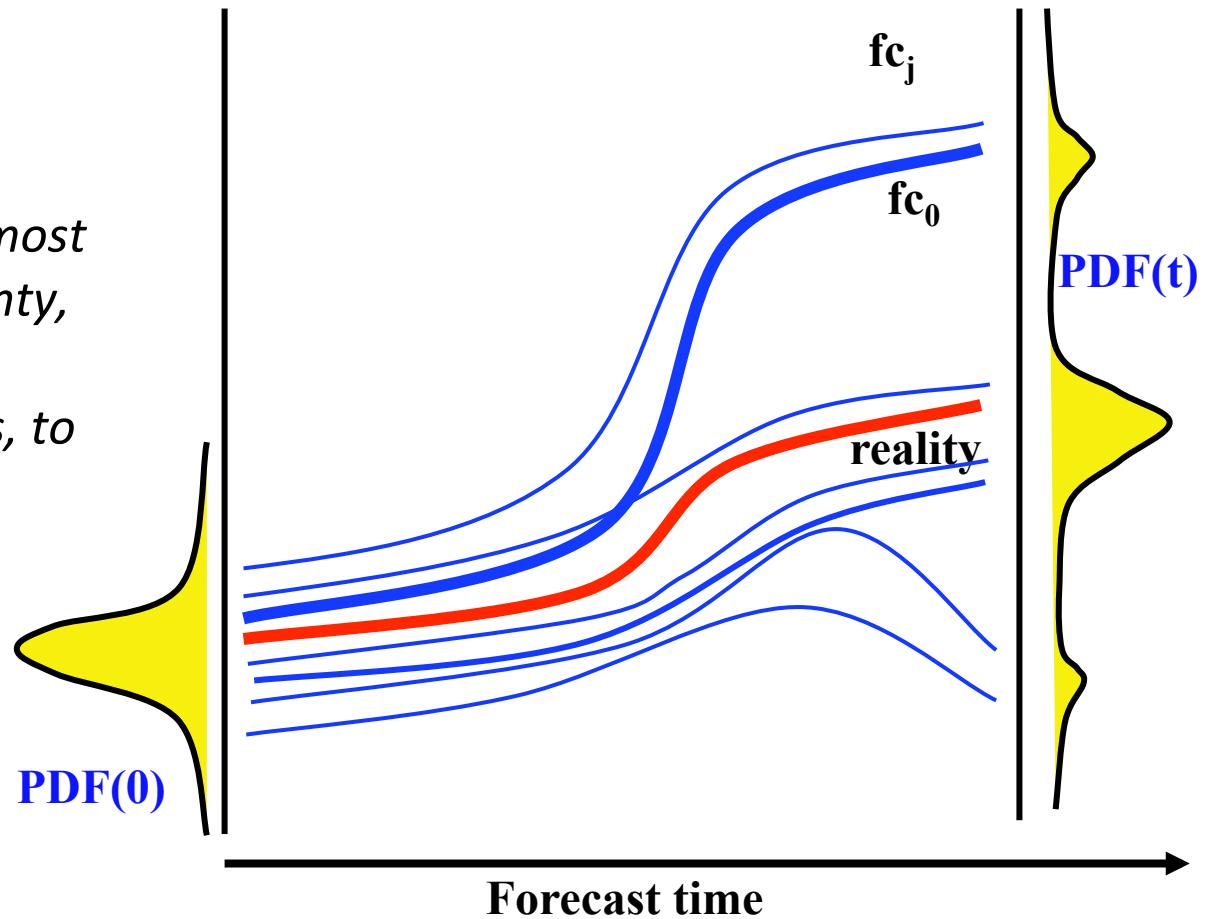
*Acknowledgement: everyone at ECMWF,
in particular people in the Research Department
and the Predictability Division.*



Ensembles: why?

To predict the time evolution of the probability density function of forecast states.

In other words, to predict the most likely scenario and its uncertainty, expressed e.g. in terms of probabilities of weather events, to estimate the forecast confidence, ..



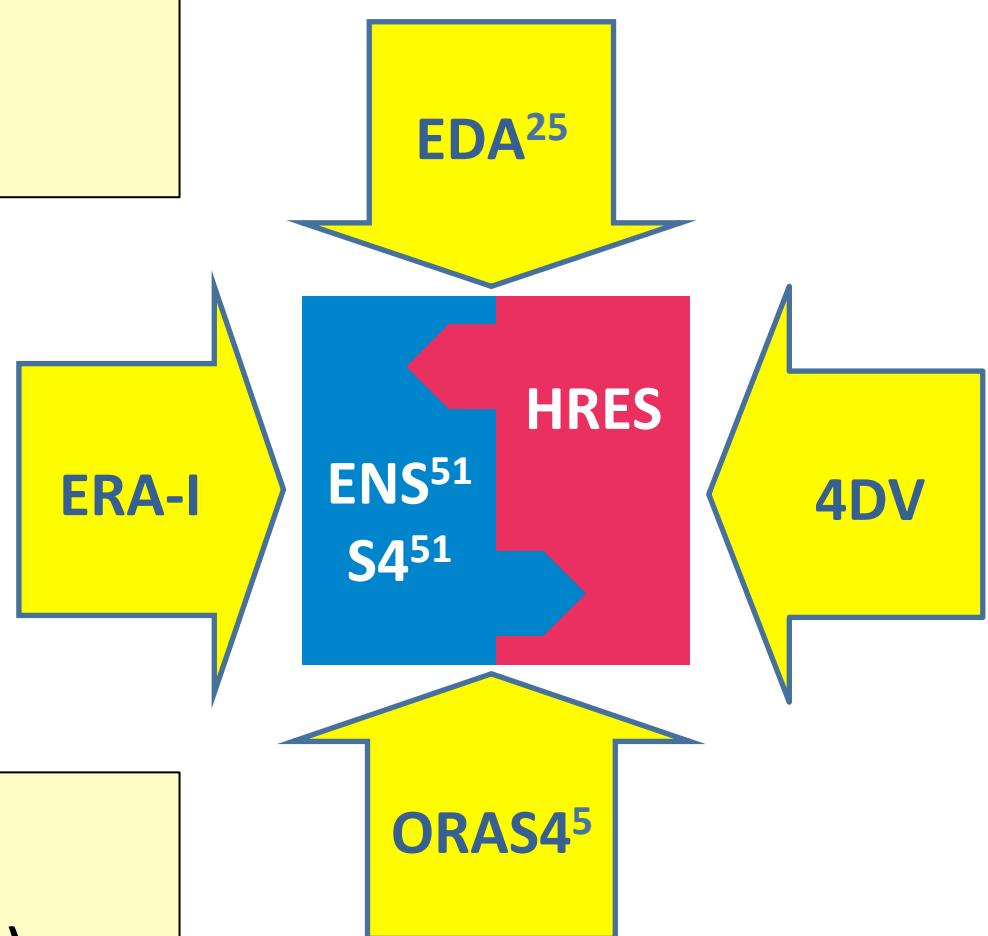


Ensembles: how do we do it at ECMWF?

$\text{PDF}(0) \ll 4\text{DV} + \text{EDA}^{25} + \text{ORAS4}^5$

$\text{PDF}(0) \ll \text{ERA-I} + \text{ORAS4}^5$ (*refc suite*)

$\text{PDF}(T) \ll \text{HRES} + \text{ENS}^{51}/\text{S4}^{51}$

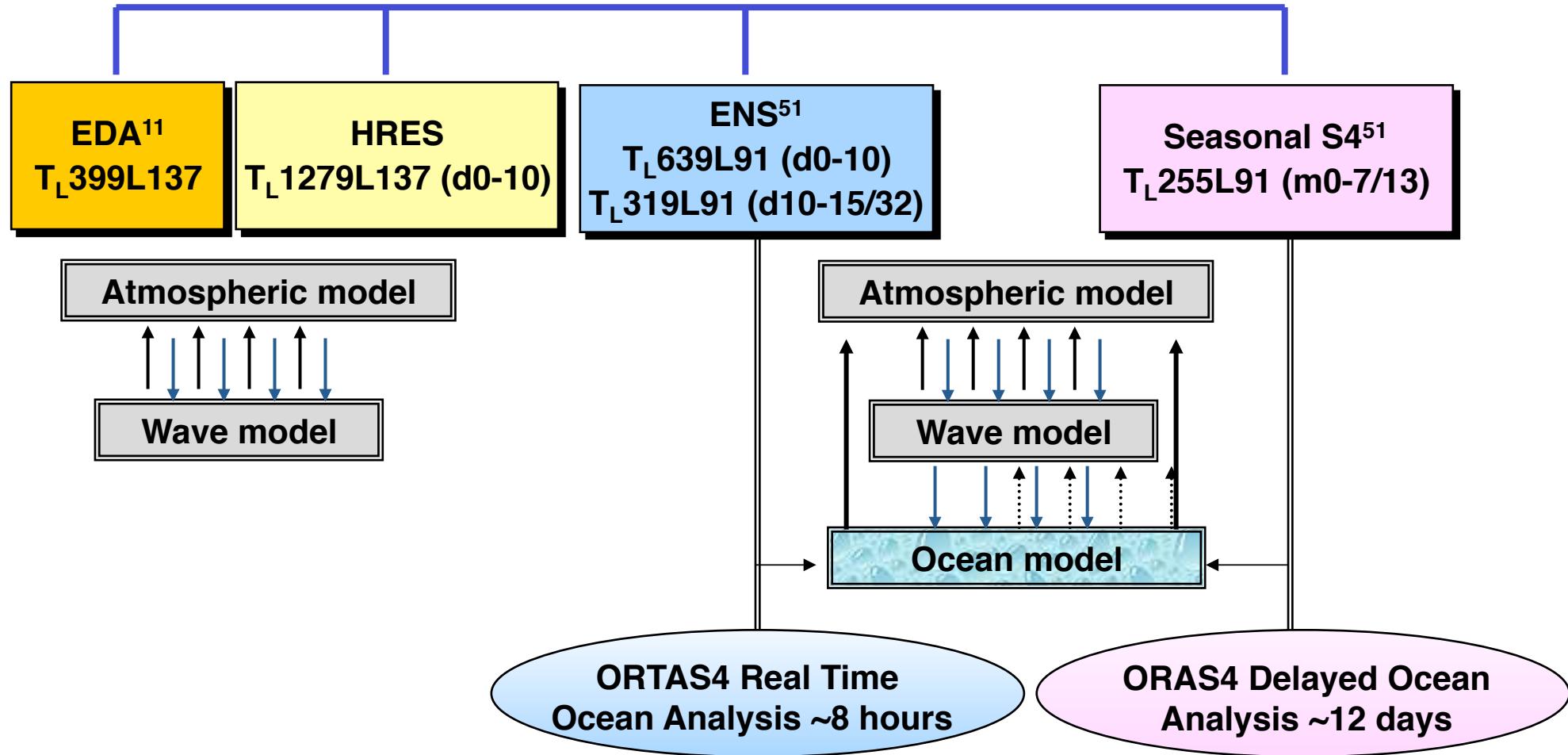


System components simulate the effect of:

- Observation uncertainties
- Model uncertainties (2 stochastic schemes)



The components of ECMWF Integrated Forecasting System





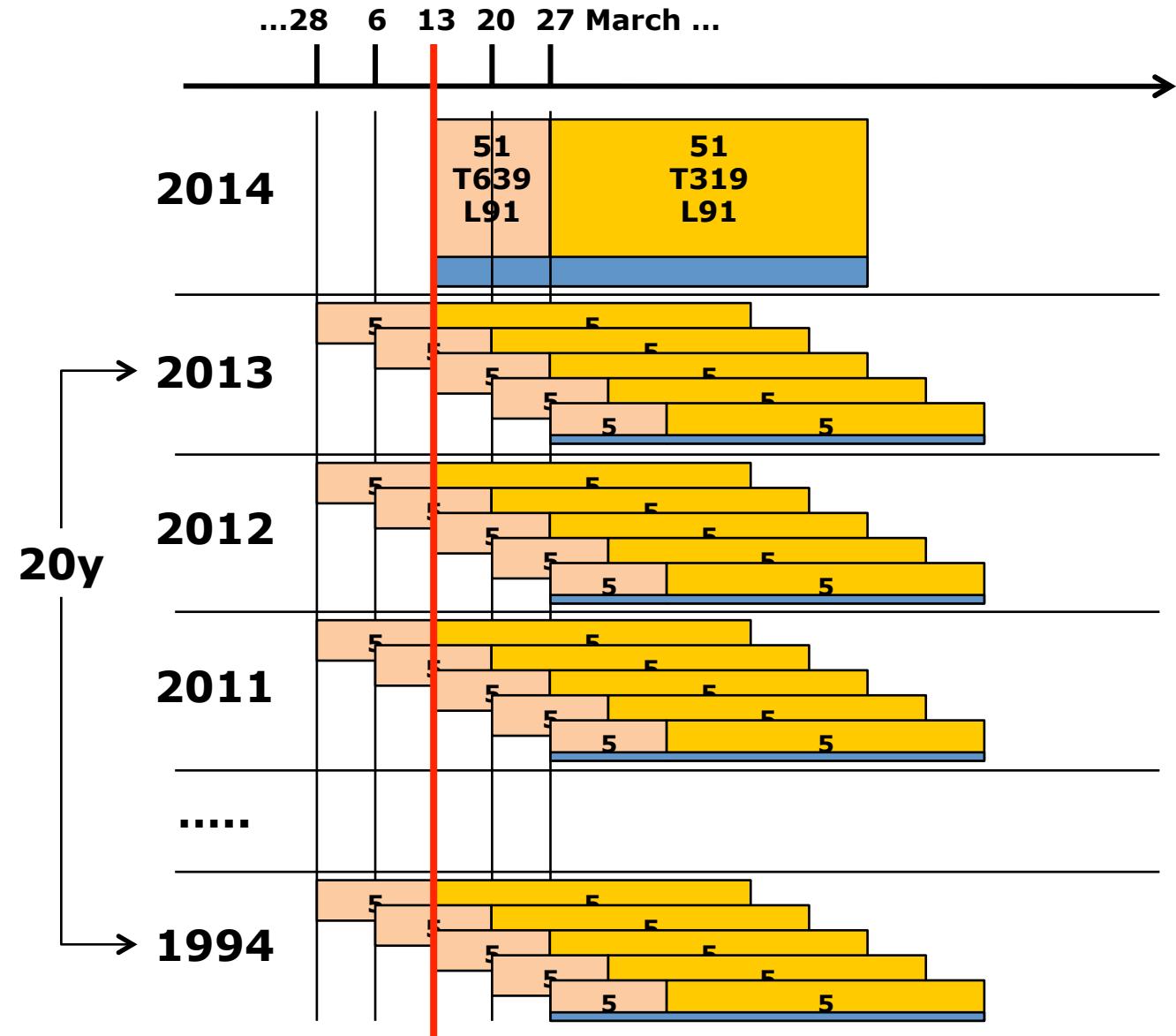
The ENS refc suite provides an estimate the M-climate

Following *Thomas Hamill* work, a re-fc suite is part of all ECMWF ensemble systems. Each day, the M-climate is estimated using

500 EPS re-forecasts:

- 20 years (1994 – 2013)
- 5 ICs (-14d,-7d,0,+7d, +14d)
- 5 members

Some of the ENS products (e.g. the Extreme Forecast Indices) are bias corrected and/or calibrated using the model climate.

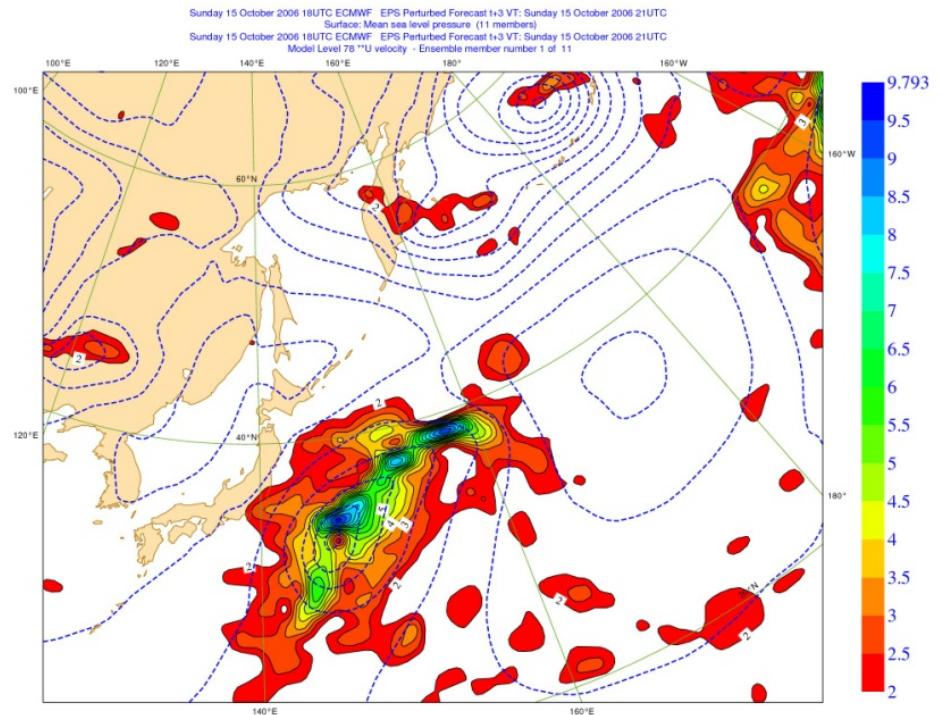
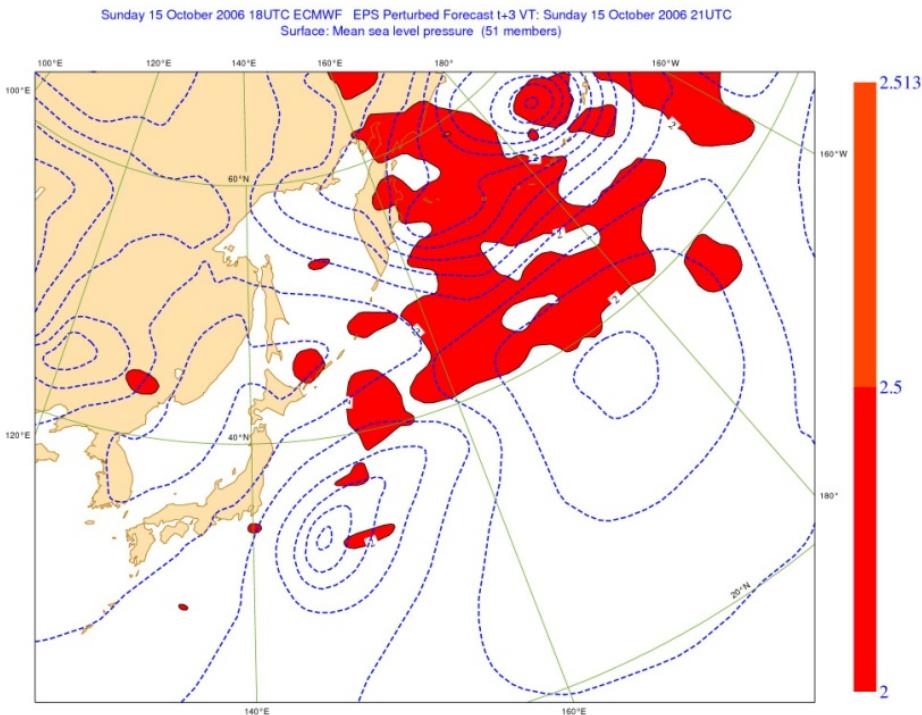




EDA provides flow-dependent analysis error var/covar

At initial time, the EDA captures the **flow-dependent** analysis uncertainty due to observation and model uncertainty.

**U850 background error standard deviation
Randomization method (left) – EDA (from cy36r4, right)**



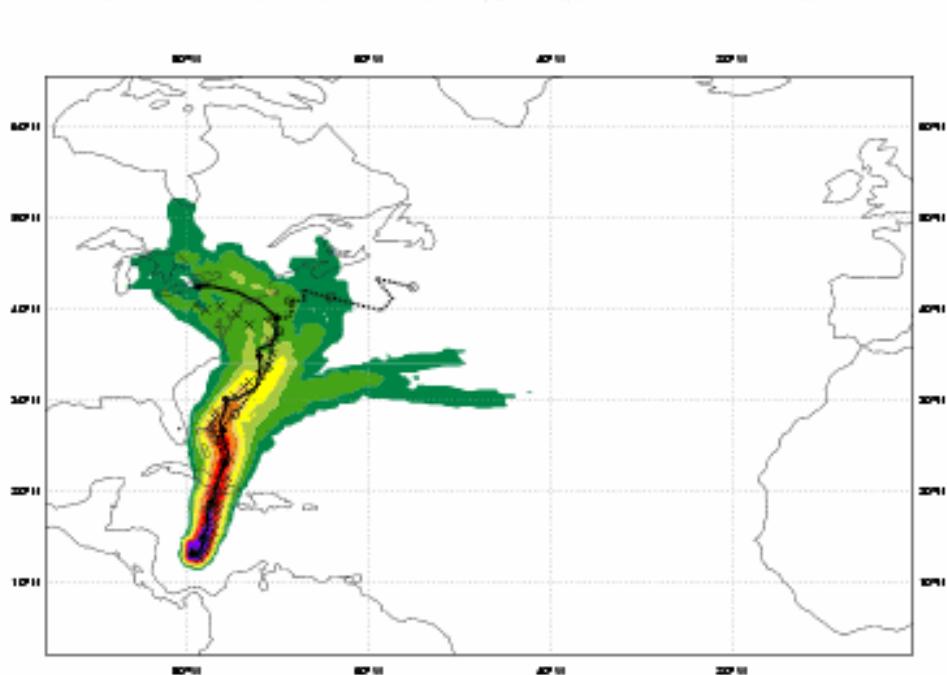
(from M Bonavita)



ENS provides probabilities of weather scenarii

Date 20121023 00 UTC @ ECMWF

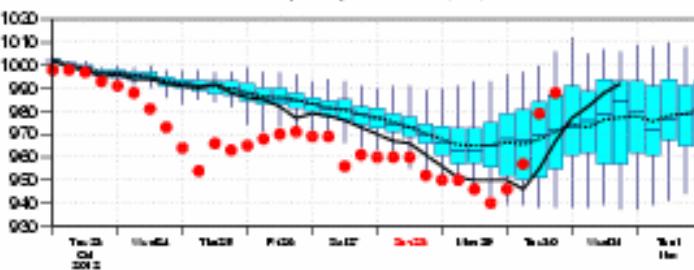
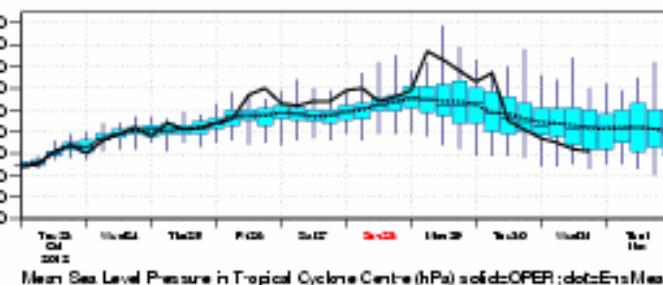
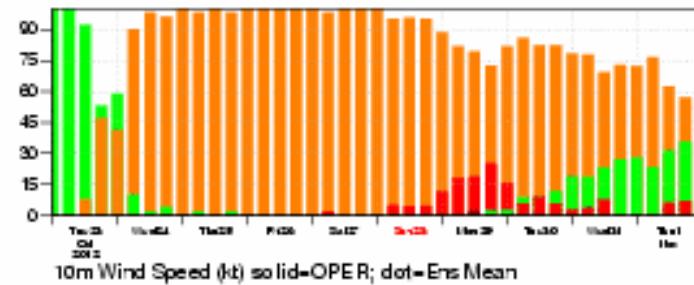
Probability that **SANDY** will pass within 120 km radius during the next 240 hours
tracks: solid=OPER; dot=Ens Mean [reported minimum central pressure (hPa) 988]



List of ensemble members numbers forecast Tropical Cyclone Intensity category in colours: TD[up to 33] TS[34-63] HR1[64-82] HR2[83-95] HR3[> 95 kt]

+000 hPa	d 01	d 02	d 03	d 04	d 05	d 06	d 07	d 08	d 09	d 10	d 11	d 12	d 13	d 14	d 15	d 16	d 17	d 18	d 19	d 20	d 21	d 22	d 23	d 24	d 25	d 26	d 27	d 28	d 29	d 30	d 31	d 32	d 33	d 34	d 35	d 36	d 37	d 38	d 39	d 40	d 41	d 42	d 43	d 44	d 45	d 46	d 47	d 48	d 49	d 50	d 51	d 52	d 53	d 54	d 55	d 56	d 57	d 58	d 59	d 60	d 61	d 62	d 63	d 64	d 65	d 66	d 67	d 68	d 69	d 70	d 71	d 72	d 73	d 74	d 75	d 76	d 77	d 78	d 79	d 80	d 81	d 82	d 83	d 84	d 85	d 86	d 87	d 88	d 89	d 90	d 91	d 92	d 93	d 94	d 95	d 96	d 97	d 98	d 99	d 100
+000 hPa	d 01	d 02	d 03	d 04	d 05	d 06	d 07	d 08	d 09	d 10	d 11	d 12	d 13	d 14	d 15	d 16	d 17	d 18	d 19	d 20	d 21	d 22	d 23	d 24	d 25	d 26	d 27	d 28	d 29	d 30	d 31	d 32	d 33	d 34	d 35	d 36	d 37	d 38	d 39	d 40	d 41	d 42	d 43	d 44	d 45	d 46	d 47	d 48	d 49	d 50	d 51	d 52	d 53	d 54	d 55	d 56	d 57	d 58	d 59	d 60	d 61	d 62	d 63	d 64	d 65	d 66	d 67	d 68	d 69	d 70	d 71	d 72	d 73	d 74	d 75	d 76	d 77	d 78	d 79	d 80	d 81	d 82	d 83	d 84	d 85	d 86	d 87	d 88	d 89	d 90	d 91	d 92	d 93	d 94	d 95	d 96	d 97	d 98	d 99	d 100
+000 hPa	d 01	d 02	d 03	d 04	d 05	d 06	d 07	d 08	d 09	d 10	d 11	d 12	d 13	d 14	d 15	d 16	d 17	d 18	d 19	d 20	d 21	d 22	d 23	d 24	d 25	d 26	d 27	d 28	d 29	d 30	d 31	d 32	d 33	d 34	d 35	d 36	d 37	d 38	d 39	d 40	d 41	d 42	d 43	d 44	d 45	d 46	d 47	d 48	d 49	d 50	d 51	d 52	d 53	d 54	d 55	d 56	d 57	d 58	d 59	d 60	d 61	d 62	d 63	d 64	d 65	d 66	d 67	d 68	d 69	d 70	d 71	d 72	d 73	d 74	d 75	d 76	d 77	d 78	d 79	d 80	d 81	d 82	d 83	d 84	d 85	d 86	d 87	d 88	d 89	d 90	d 91	d 92	d 93	d 94	d 95	d 96	d 97	d 98	d 99	d 100
+000 hPa	d 01	d 02	d 03	d 04	d 05	d 06	d 07	d 08	d 09	d 10	d 11	d 12	d 13	d 14	d 15	d 16	d 17	d 18	d 19	d 20	d 21	d 22	d 23	d 24	d 25	d 26	d 27	d 28	d 29	d 30	d 31	d 32	d 33	d 34	d 35	d 36	d 37	d 38	d 39	d 40	d 41	d 42	d 43	d 44	d 45	d 46	d 47	d 48	d 49	d 50	d 51	d 52	d 53	d 54	d 55	d 56	d 57	d 58	d 59	d 60	d 61	d 62	d 63	d 64	d 65	d 66	d 67	d 68	d 69	d 70	d 71	d 72	d 73	d 74	d 75	d 76	d 77	d 78	d 79	d 80	d 81	d 82	d 83	d 84	d 85	d 86	d 87	d 88	d 89	d 90	d 91	d 92	d 93	d 94	d 95	d 96	d 97	d 98	d 99	d 100
+000 hPa	d 01	d 02	d 03	d 04	d 05	d 06	d 07	d 08	d 09	d 10	d 11	d 12	d 13	d 14	d 15	d 16	d 17	d 18	d 19	d 20	d 21	d 22	d 23	d 24	d 25	d 26	d 27	d 28	d 29	d 30	d 31	d 32	d 33	d 34	d 35	d 36	d 37	d 38	d 39	d 40	d 41	d 42	d 43	d 44	d 45	d 46	d 47	d 48	d 49	d 50	d 51	d 52	d 53	d 54	d 55	d 56	d 57	d 58	d 59	d 60	d 61	d 62	d 63	d 64	d 65	d 66	d 67	d 68	d 69	d 70	d 71	d 72	d 73	d 74	d 75	d 76	d 77	d 78	d 79	d 80	d 81	d 82	d 83	d 84	d 85	d 86	d 87	d 88	d 89	d 90	d 91	d 92	d 93	d 94	d 95	d 96	d 97	d 98	d 99	d 100

Probability (%) of Tropical Cyclone Intensity falling in each category
TD[up to 33] TS[34-63] HR1[64-82] HR2[83-95] HR3[> 95 kt]





ENS predicts the probability of ‘rare’ weather

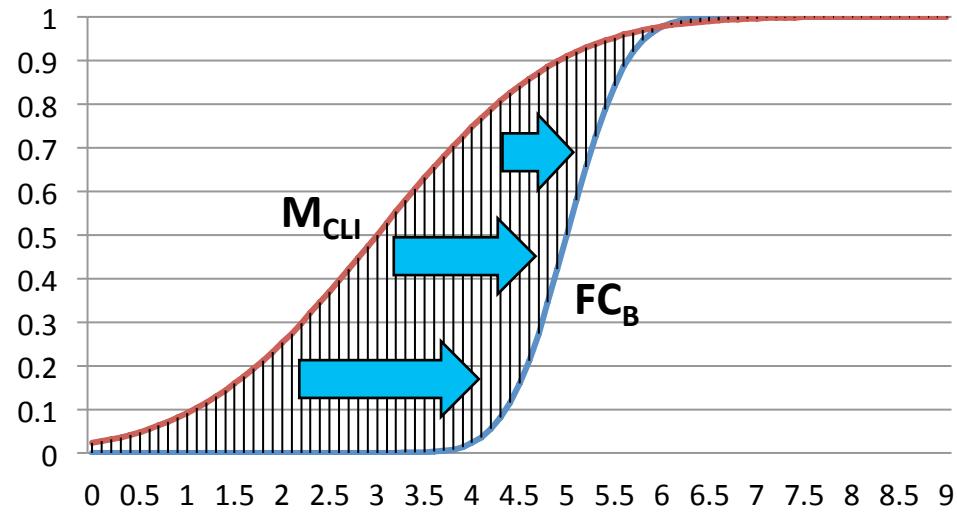
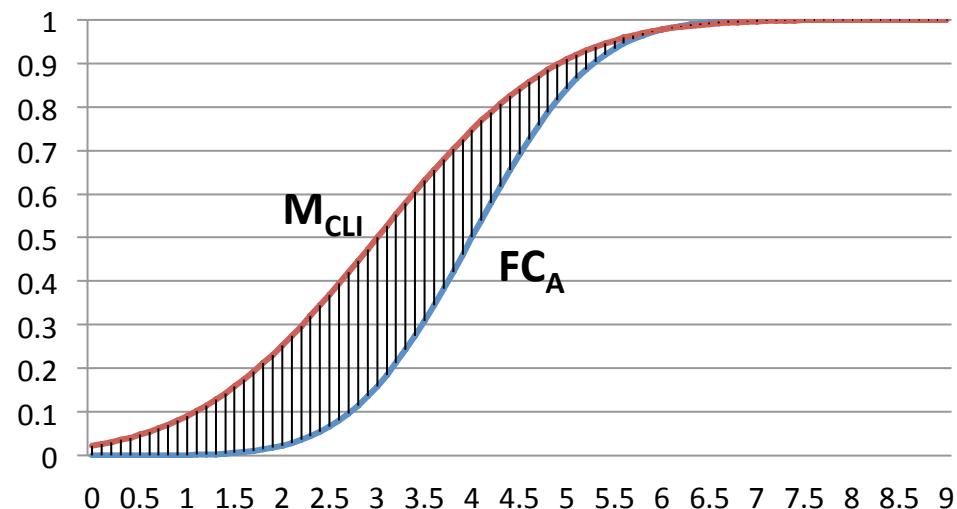
By comparing the model climate CDF^{MCLI} and the forecast CDFs (coloured lines), we can define the Extreme Forecast Index (EFI) and ***predict extremes as seen by the model.***

The EFI is the average difference between the CDF^{MCLI} and the fc CDF:

$$EFI = \frac{2}{\pi} \int_0^1 \frac{p - CDF(p)}{\sqrt{p(1-p)}} dp$$

The figures shows the following fc/obs PDFs:

- $MCLI=N(3.0,1.5)$
- $FC_A=N(4.0,1.0) \rightarrow EFI_A=+40\%$
- $FC_B=N(5.0,0.5) \rightarrow EFI_B=+59\%$

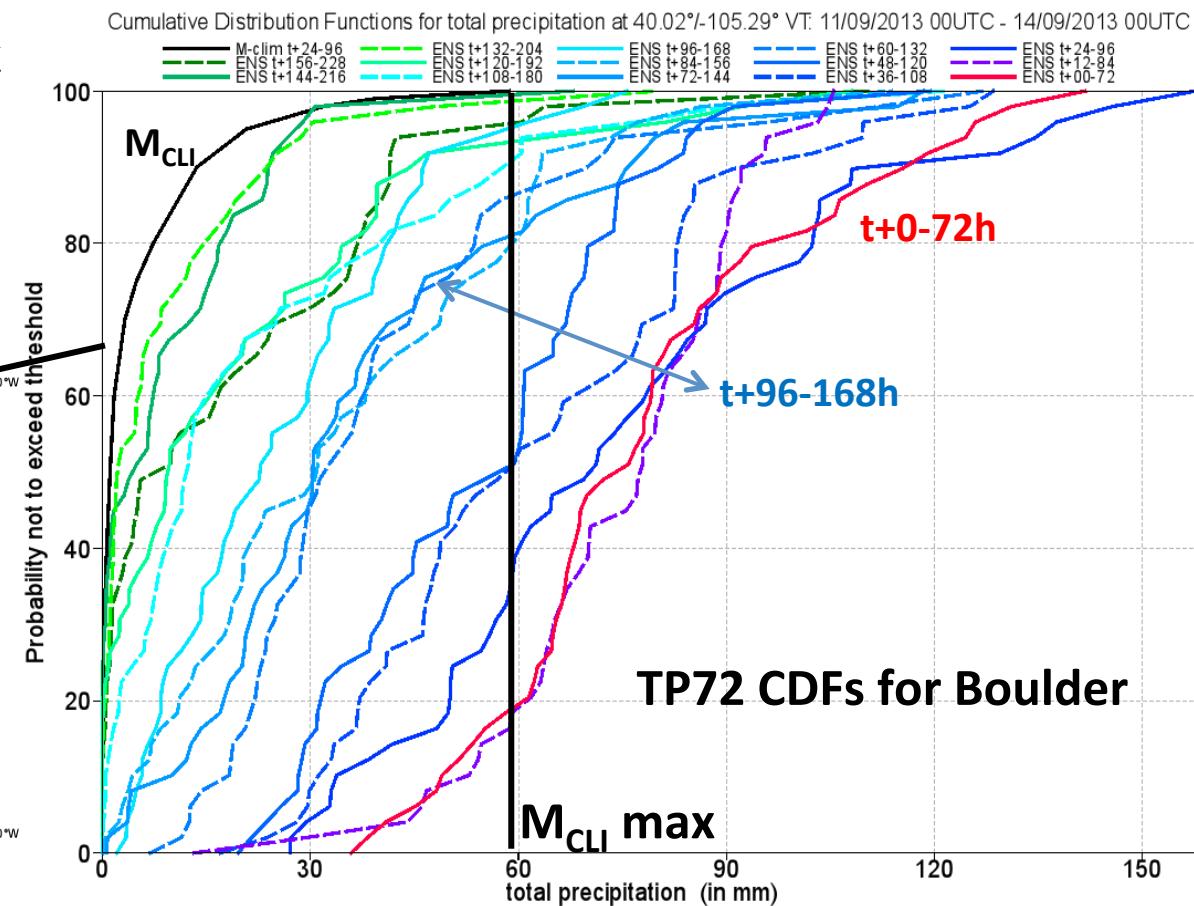
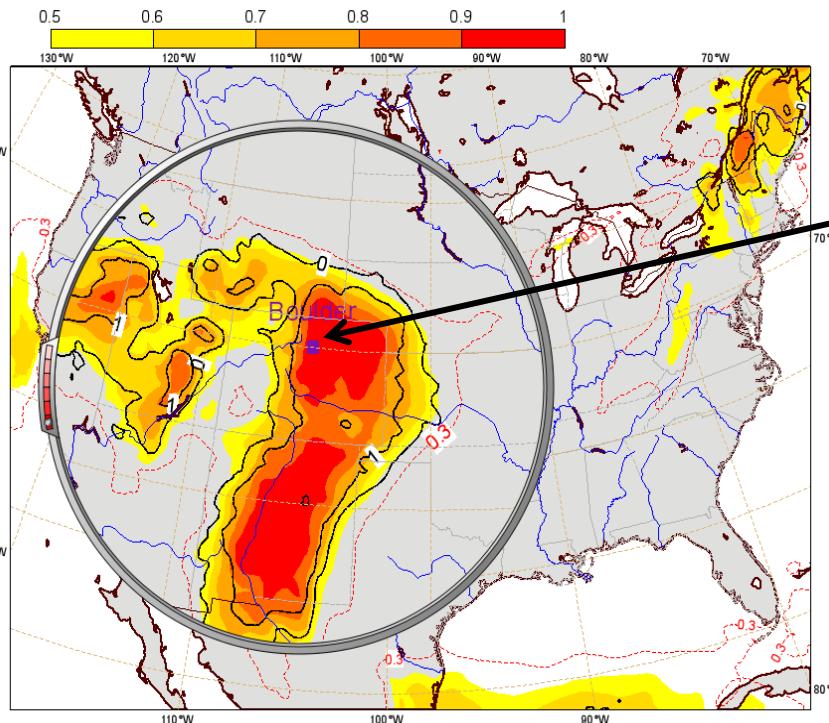




Boulder, ENS EFI for TP72 fcs vt 11@00UTC-14@00UTC

ENS t+0-72h Extreme Forecast Index

Wed 11 Sep 2013 00UTC @ECMWF VT: Wed 11 Sep 2013 00UTC - Sat 14 Sep 2013 00UTC 0-72h
Extreme forecast index and Shift of Tails (black contours 0,1,5,10,15) for: total precipitation

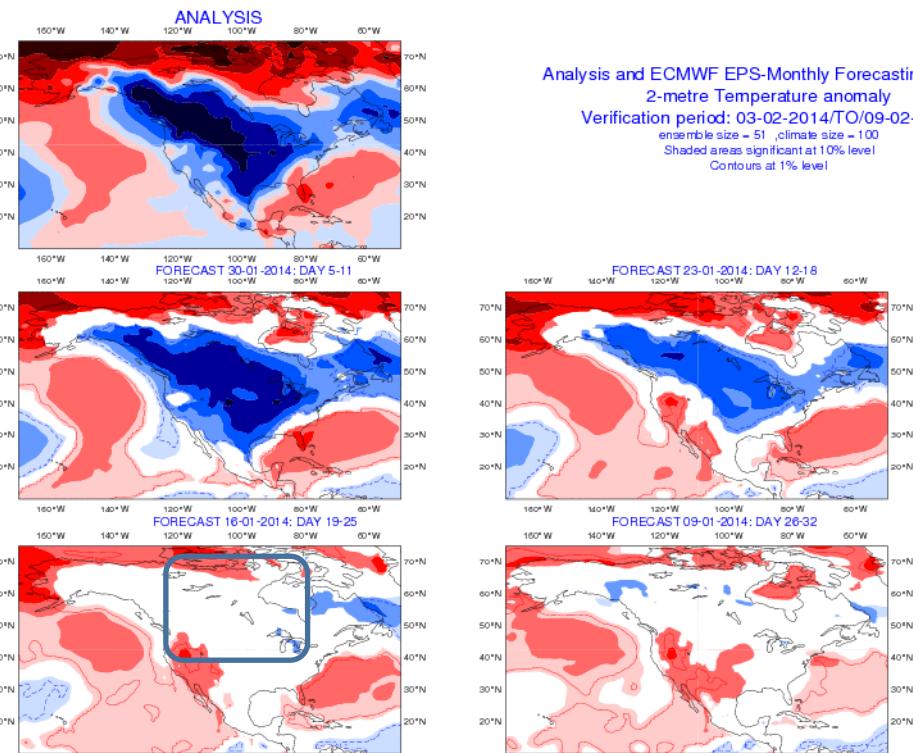


The EFI reached high values (approaching 1) in the last few forecasts preceding the onset of the event, with ~10% of the ENS members predicting rainfall beyond the climate extreme (99th percentile of the model climate).

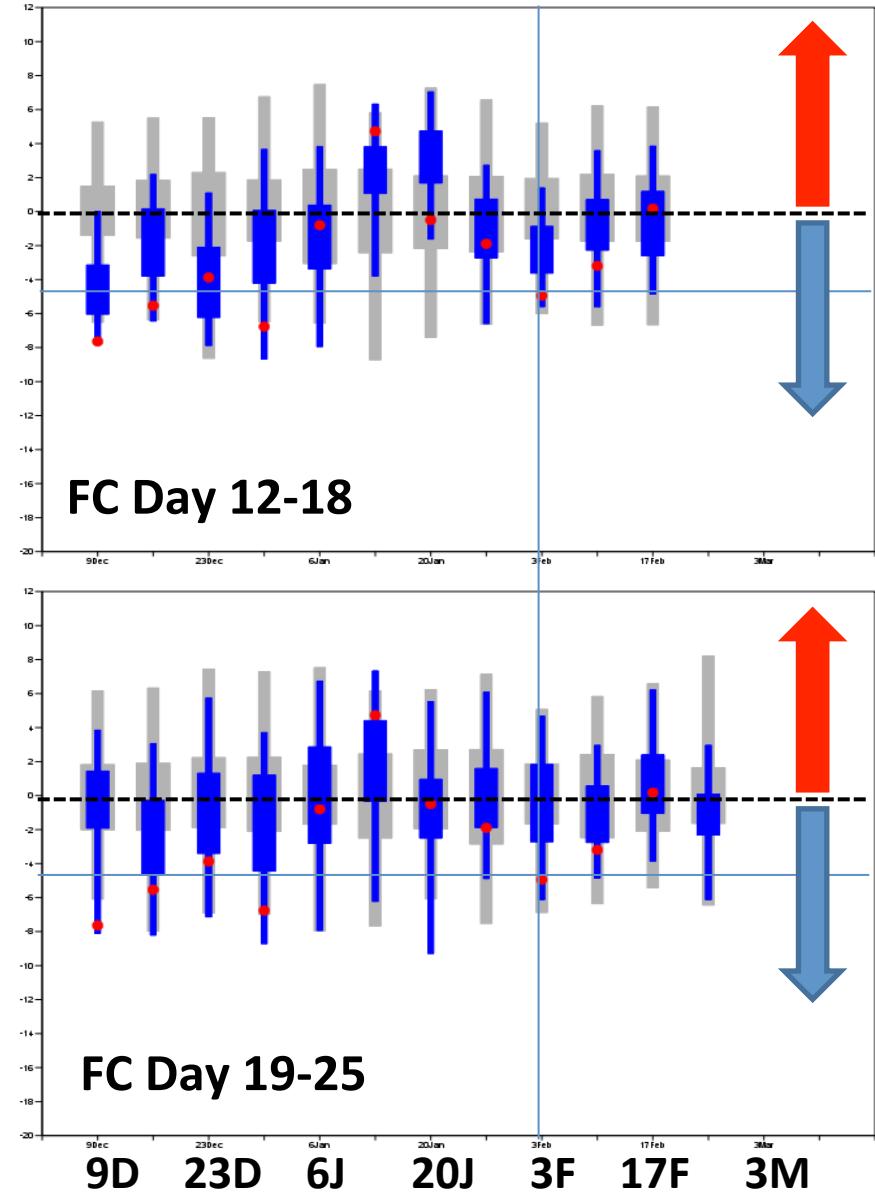
(From Ivan Tsonevski)



ENS extended to 32d predicts anomalies weeks ahead



This plot shows ENS weekly anomaly forecasts over North America in D13-JFM13 (the right panels show average 2mT over (120-60W; 40-70N)).

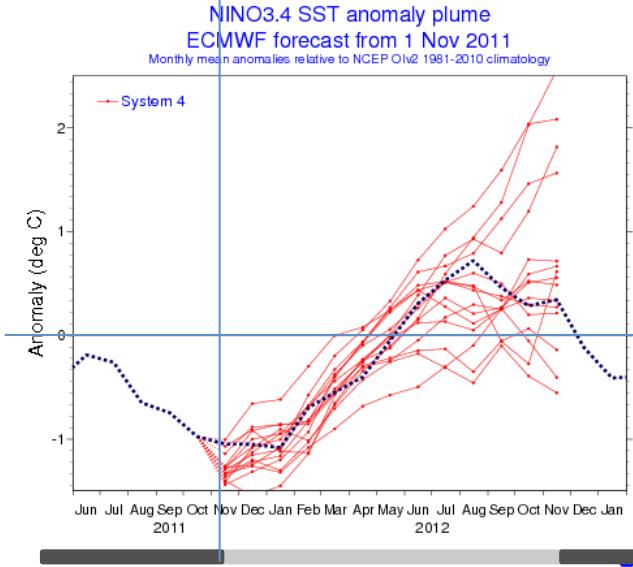




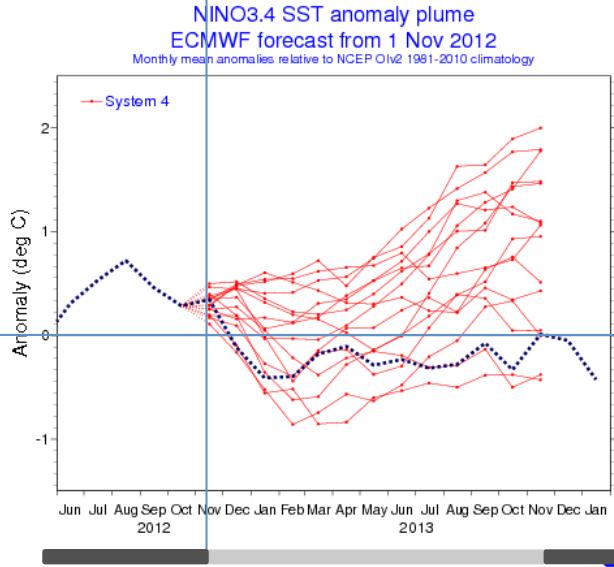
The seasonal ensemble S4 provides probabilities up to 1y

The tropics remain the area where seasonal prediction has the highest skill, as indicated e.g. by the accuracy of 1-year forecasts of SST anomaly in the Nino3.4 area.

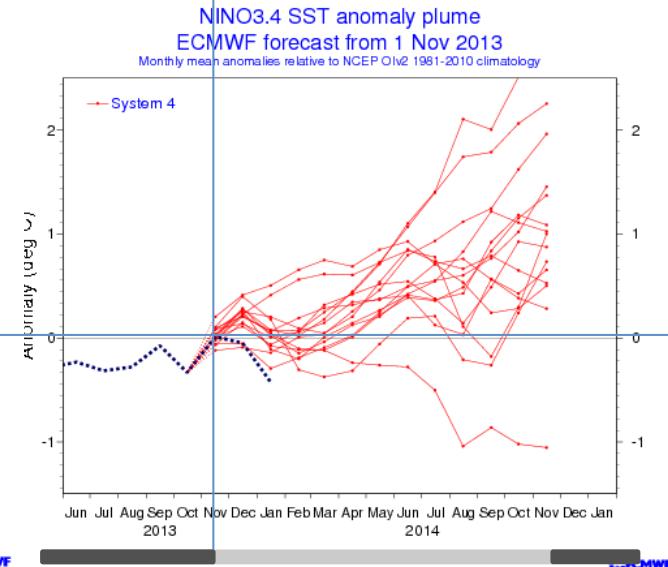
1 Nov '11 > Nov '12



1 Nov '12 > Nov '13



1 Nov '13 > Nov '14



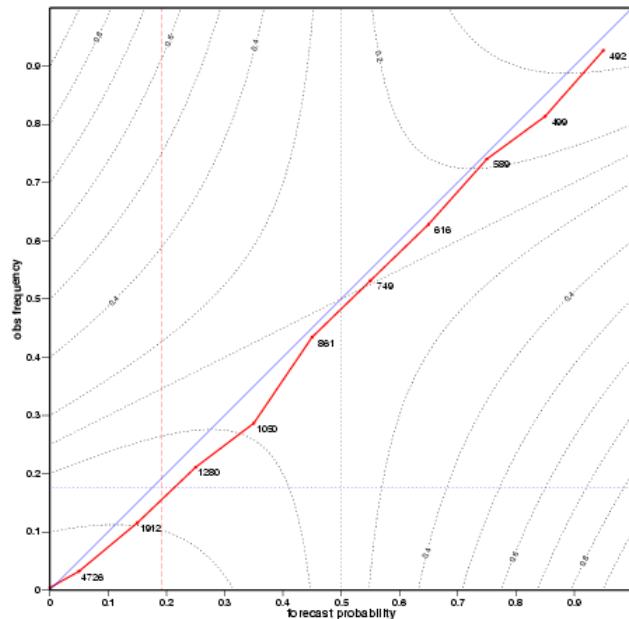


How do our ensembles perform? They must be reliable

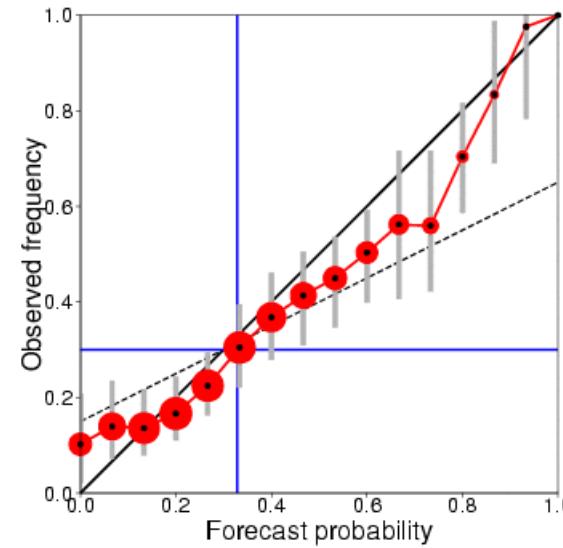
Reliability is a key property that probabilistic forecasts must have.

Reliability: ENS T850 +6d fcs over Europe

Oct12-Dec12 t + 144 Europe an T850 anomaly It -4 K
BrSc = 0.080 LCBrSkSc = 0.45 Uncertainty = 0.145



Reliability diagram for ECMWF with 15 ensemble members
Near-surface air temperature anomalies below the lower tercile
Accumulated over Europe (land and sea points)
Hindcast period 1981-2010 with start in May average over months 2 to 4
Skill scores and 95% conf. intervals (1000 samples)
Brier skill score: 0.108 (0.009, 0.183)
Reliability skill score: 0.980 (0.921, 0.991)
Resolution skill score: 0.128 (0.072, 0.203)



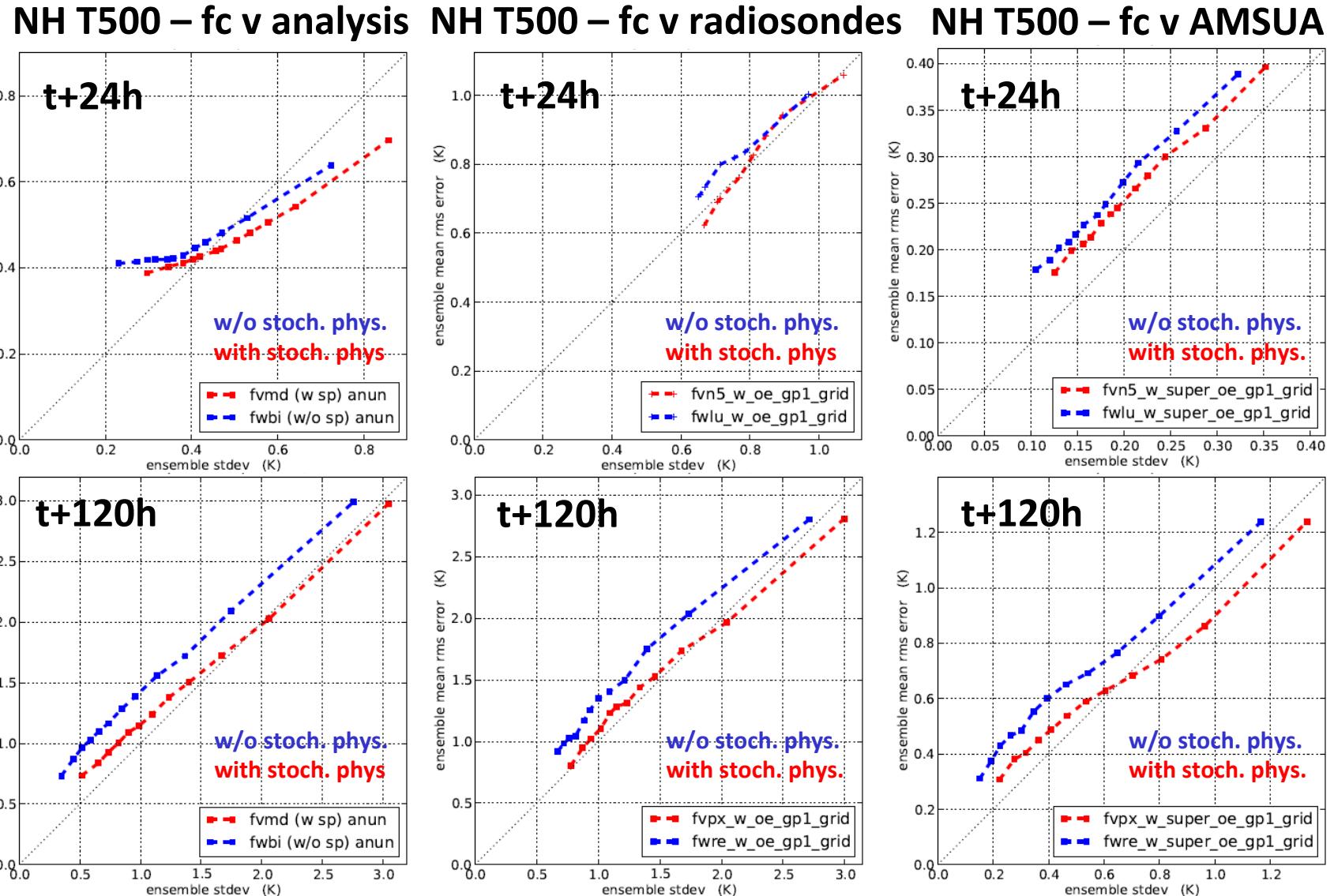
Reliability: S4 2mT 1May +2m for JJA anomaly fcs



How do they perform? They must be reliable

ENS

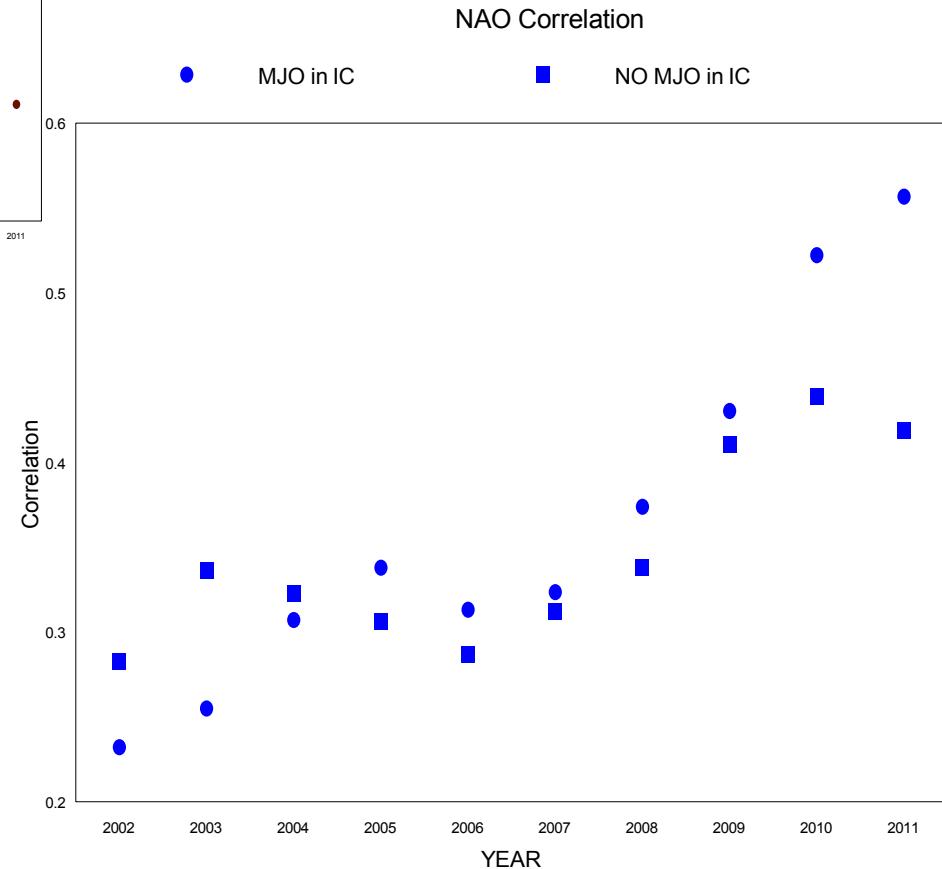
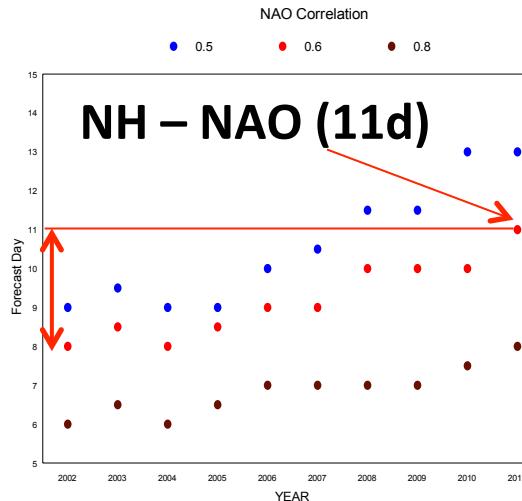
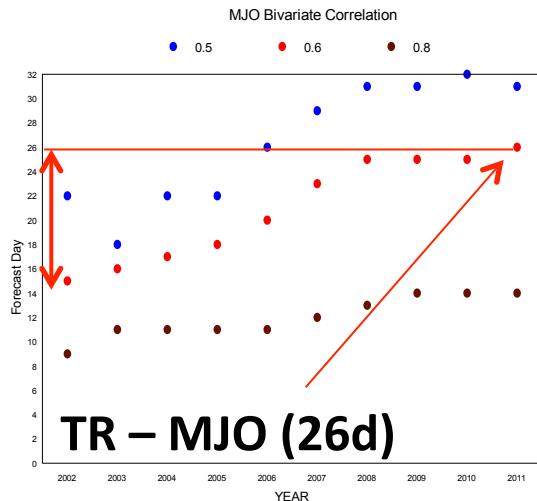
reliability is
sensitive to
verification
field
(analysis or
obs).



(From M Yamaguchi)



How do our ensembles perform? ENS MJO & NAO



The skill of monthly forecasts have been continuously improving both in the tropics for the MJO (top left) and the extra-tropics for the NAO (top right).

Improvements in the physics have led to better teleconnection between tropics and extra-tropics (bottom).

(From F Vitart)

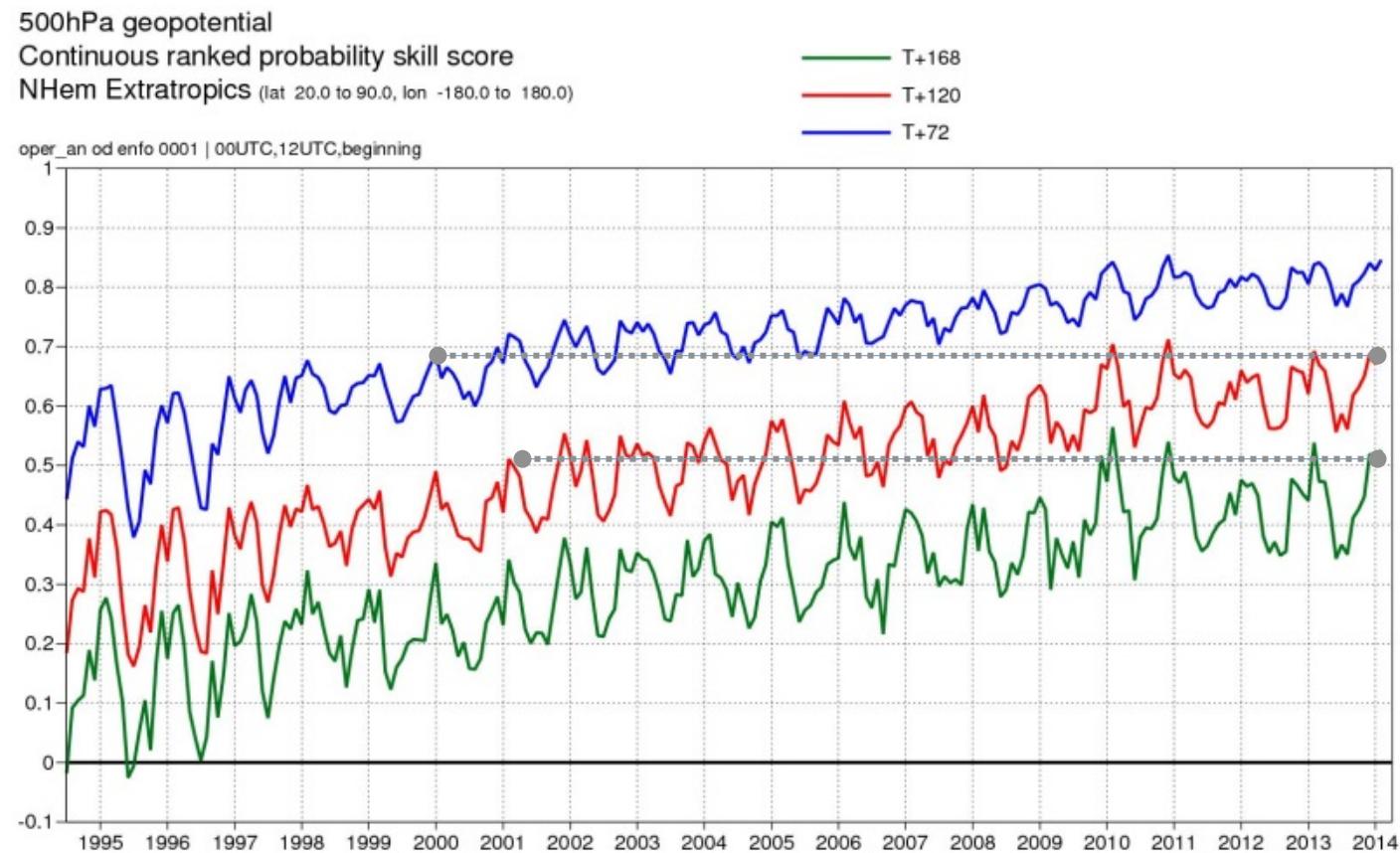


How do ensembles perform? Scores' trends CRPSS Z500 NH

Scores' trends can be used to monitor progress.

This plot shows that for upper-level fields over the NH extra-tropics , performance has been improving continuously.

Results indicate predictability gains of 1.5-2.0 days per decade.

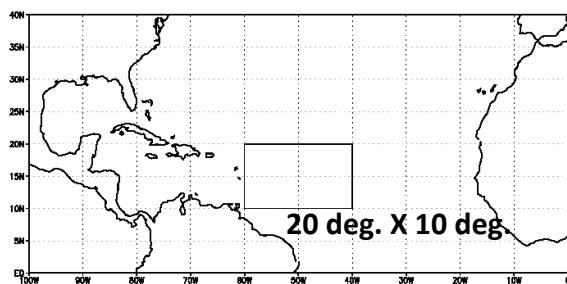
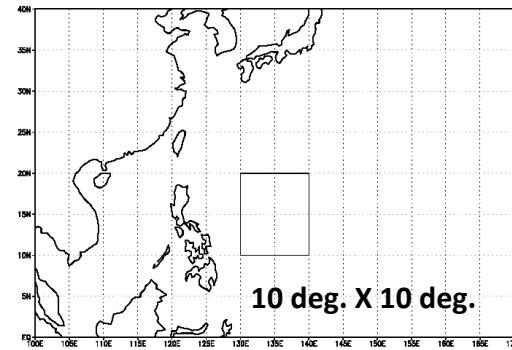




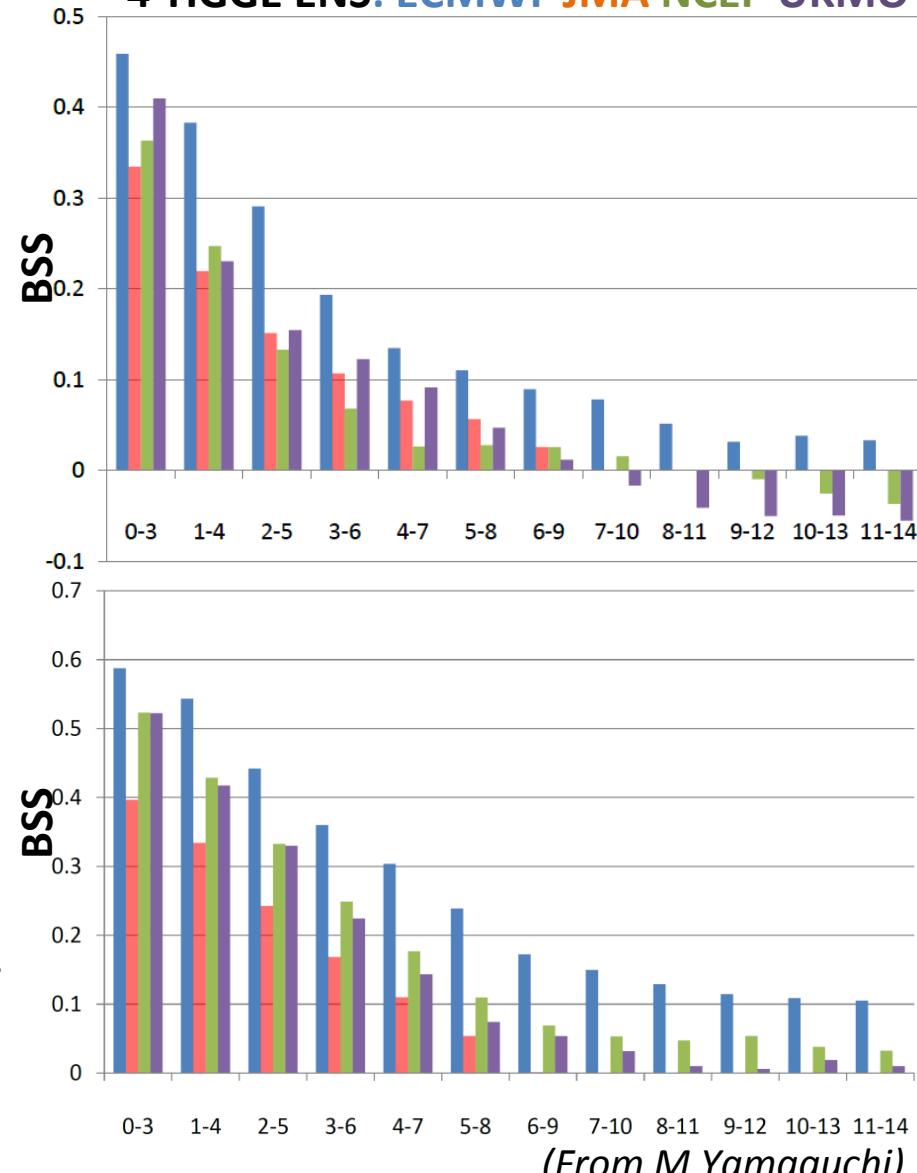
How do they perform? Hurricane strike probabilities

TIGGE-based results for Jul-Oct 2012. TC tracking data have been created with the ECMWF vortex tracker (Vitart et al. 1997, JC). Verified TCs are storms with a 'Tropical Storm' intensity or stronger (>35 knots).

Verification Areas



4 TIGGE ENS: ECMWF JMA NCEP UKMO



(From M Yamaguchi)



How do the ensembles perform? TIGGE ENS CRPSS(NH)

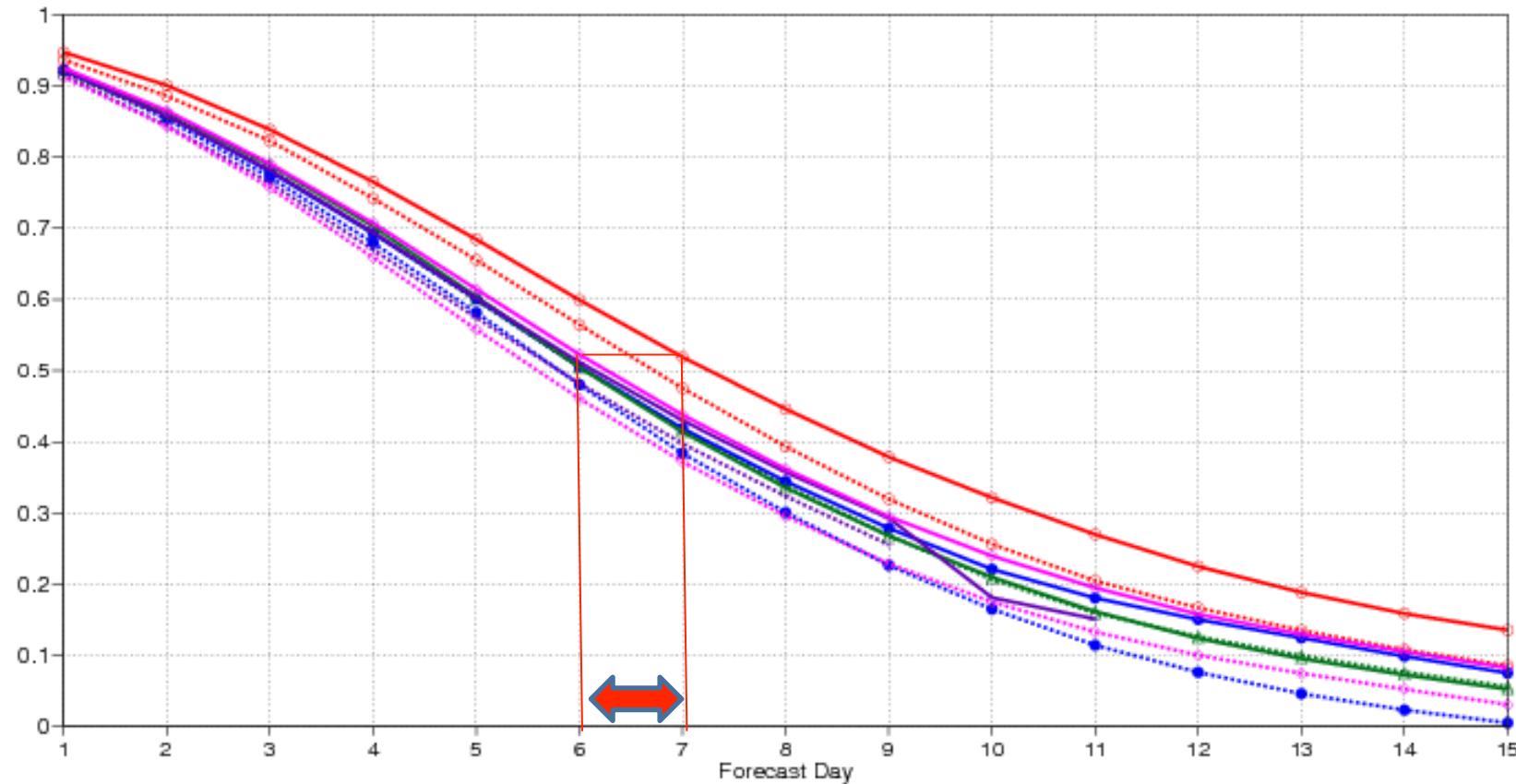
500hPa geopotential

Continuous ranked probability skill score

NHem Extratropics (Lat 20.0 to 90.0, Lon -180.0 to 180.0)

DecJanFeb

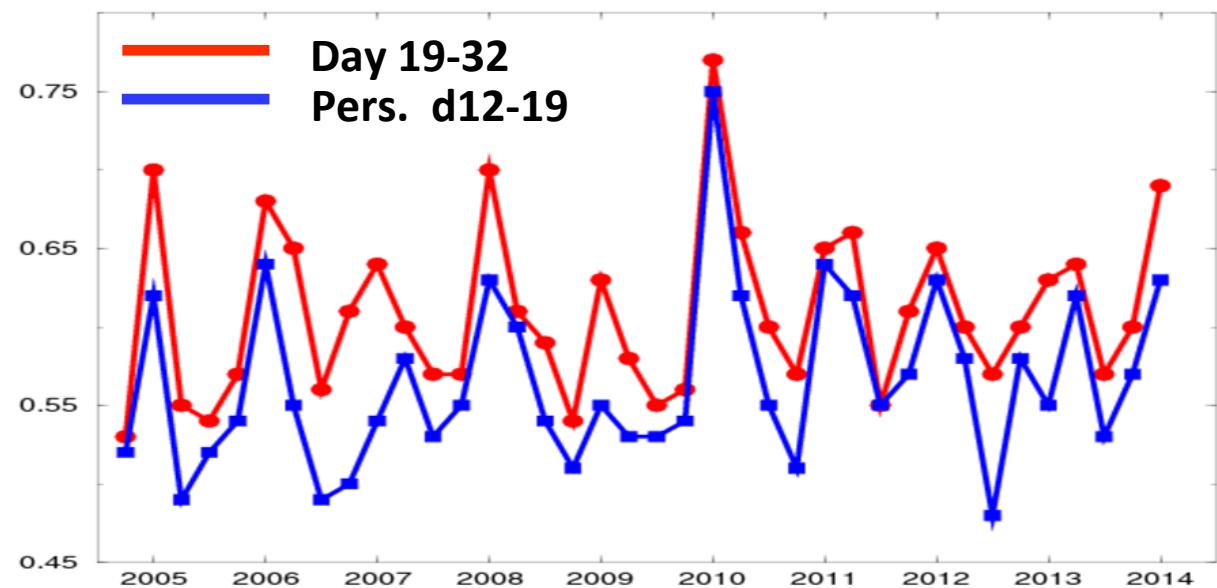
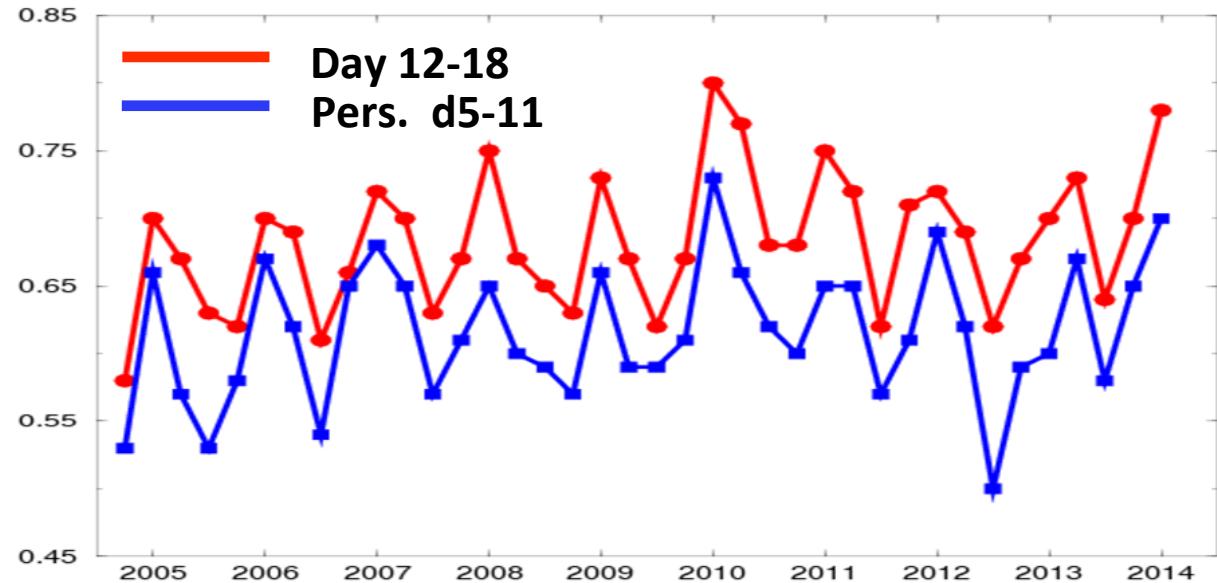
- DJF2013 CMC
- DJF2013 JMA
- DJF2013 NCEP
- DJF2013 UKMO
- DJF2013 ECMWF
- DJF2014 CMC
- DJF2014 JMA
- DJF2014 NCEP
- DJF2014 UKMO
- DJF2014 ECMWF





How do the ensembles perform? ROCA[PR(2mT>U3)] NH

ROC score for prob. Fc of 2-meter temperature in the upper tercile.





How do our ensembles perform? ENS v benchmark

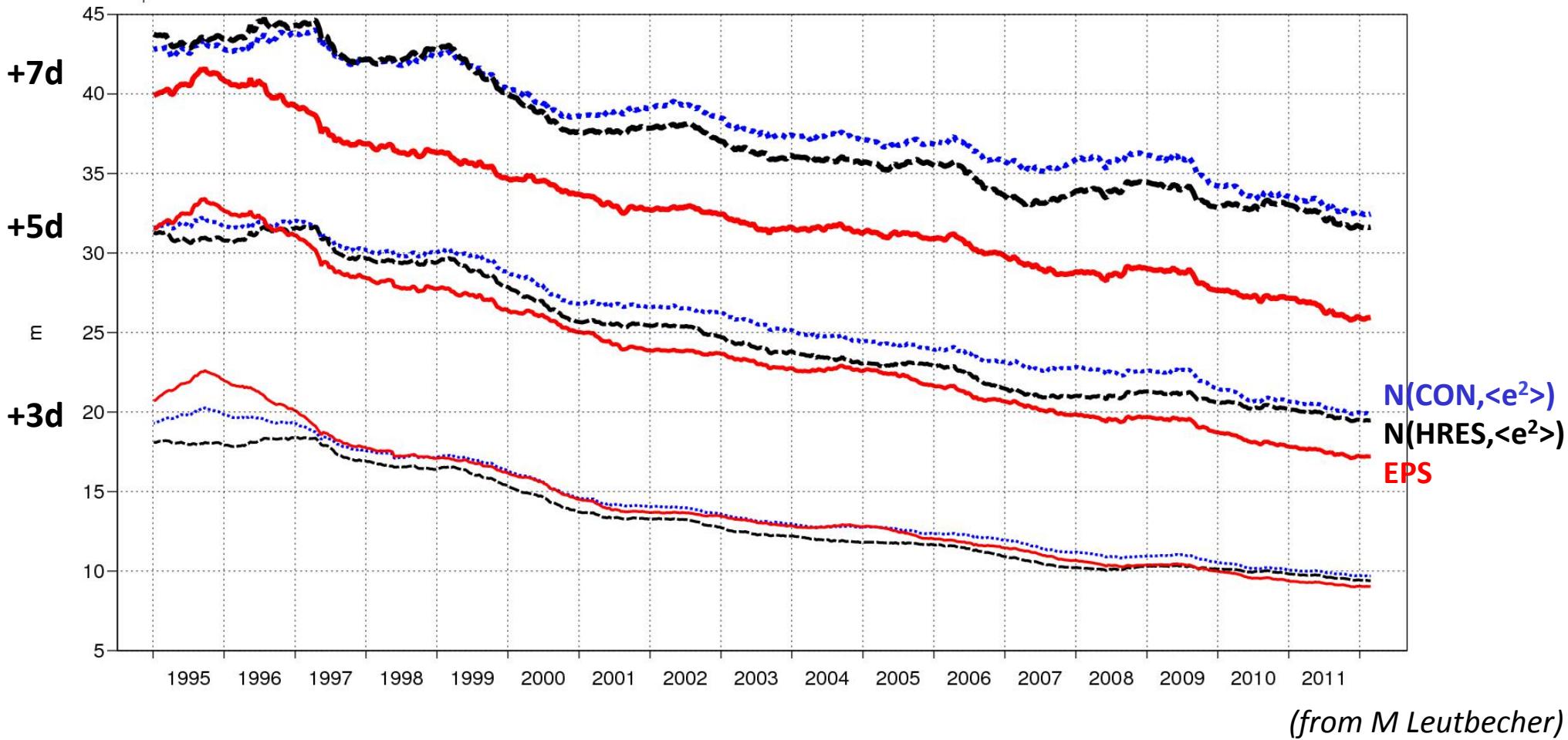
500hPa geopotential

Continuous ranked probability score

NHem Extratropics (lat 20.0 to 90.0, lon -180.0 to 180.0)

- EPS
- - - p(HR)
- p(Control)

od 0001 | 12UTC





How can we keep improving the system?

Work is progressing in several areas to further improve the forecast reliability:

- a) Improvements in the estimation of the initial conditions and in the simulation of the initial uncertainties
- b) Model improvements (e.g. moist processes, radiation, land-surface, stratosphere, ocean, ..) that include more physically-based simulation of model uncertainties
- c) System configurations (resolution, membership, update frequency, ..)

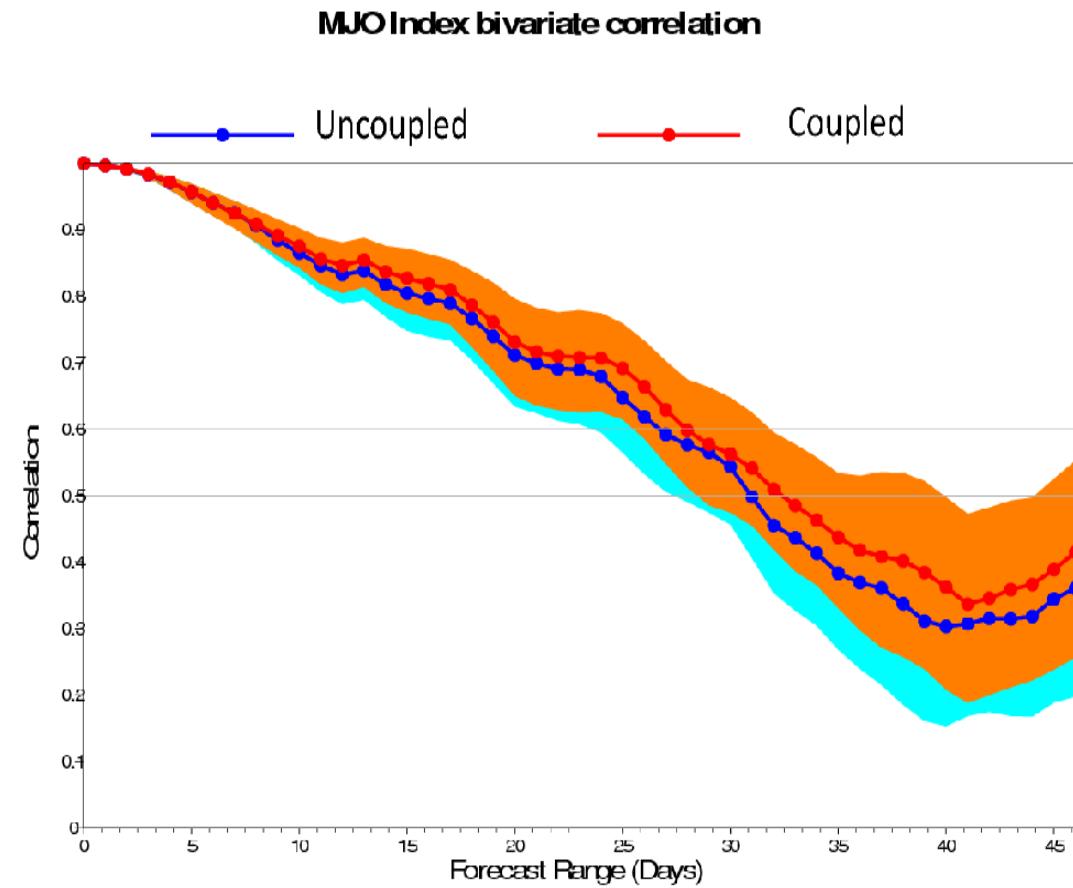


Nov 2013: ENS coupling from d0 to a better marine model ...

Nov 2013: **coupling from initial-time** to a new version of the ocean model (NEMO), with 1-way wave-currents coupling, improved skill, especially in the monthly time-range.

Work is progressing to introduce a better, unified wave-currents-sea-ice model (LIM). The new model based on NEMO is under testing at higher resolution, ORCA_025_Z75.

In 2014 we will complete a $\frac{1}{4}$ degree NEMOVAR re-analysis, and we will then implement the $\frac{1}{4}$ model first in ENS, and then in the seasonal system (S5, planned for 2016/17).

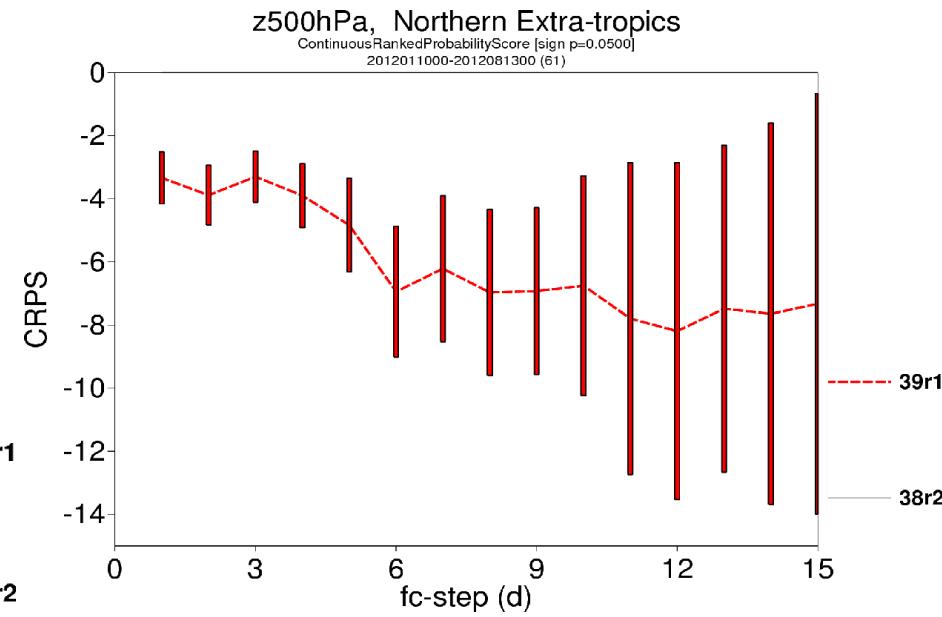
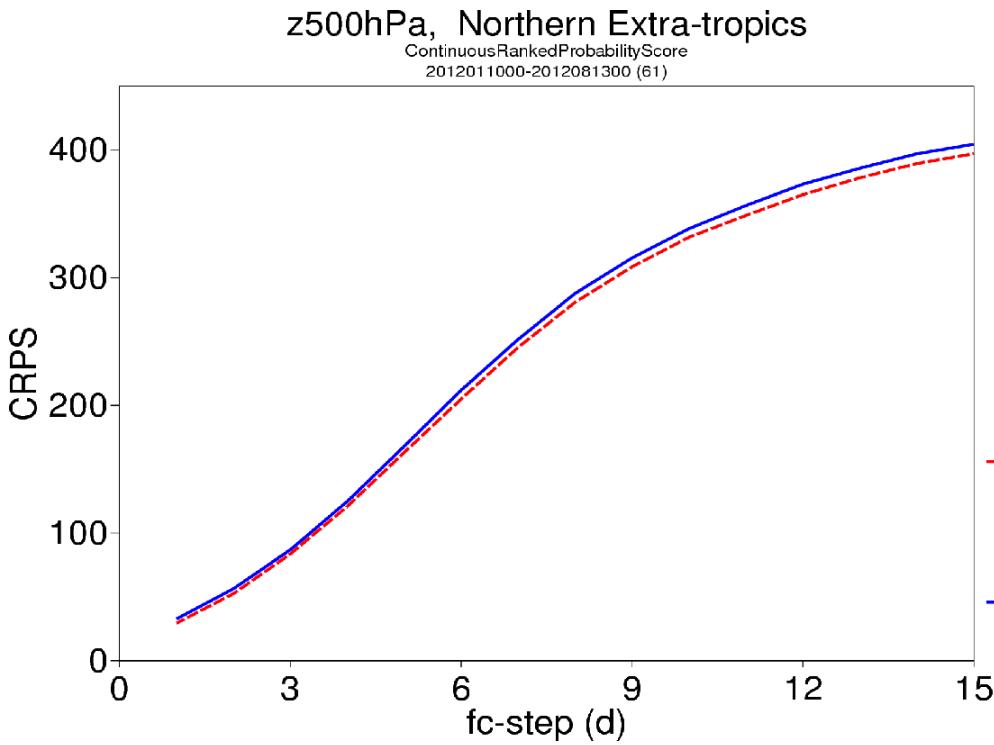
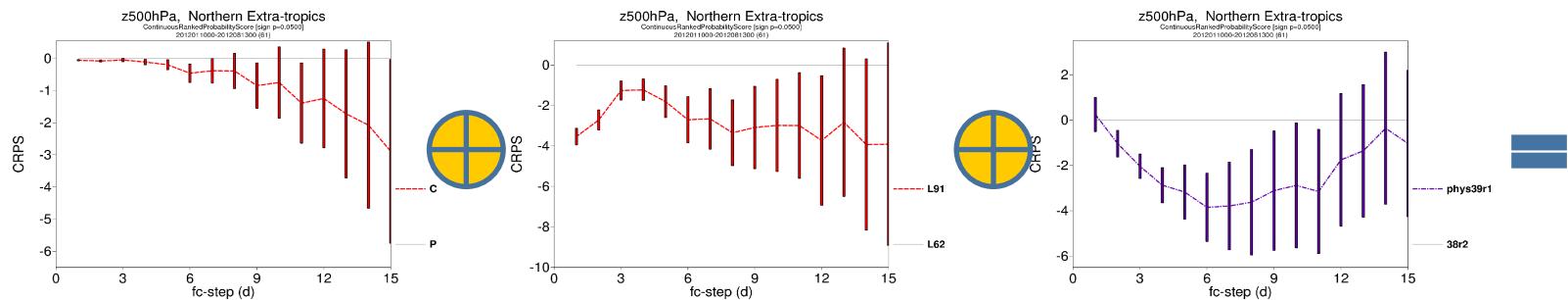


(From F Vitart)



Nov 2013: ENS coupling d0, L62>L91, new physics+SKEB

Results based on
61 cases (JFM12,
JJA12); 38r2
analyses and EDA
perturbations.

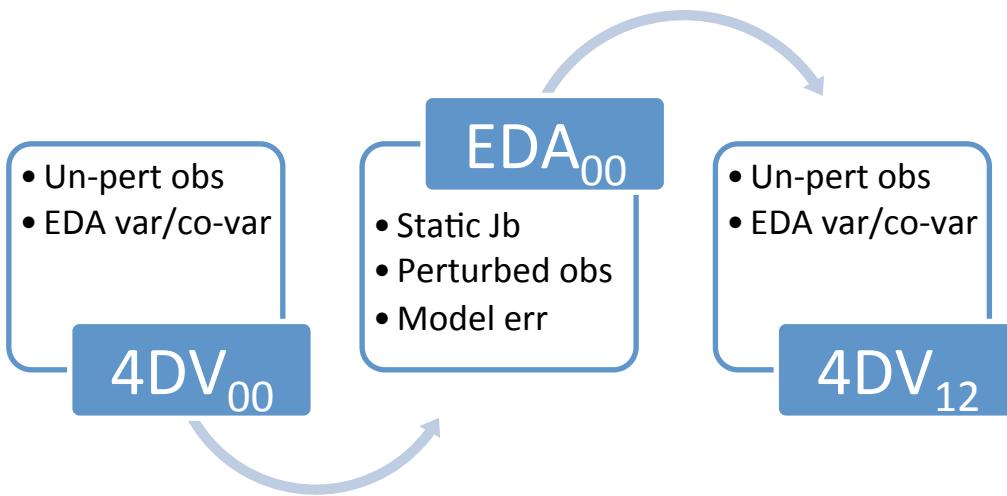


(From M Leutbecher)

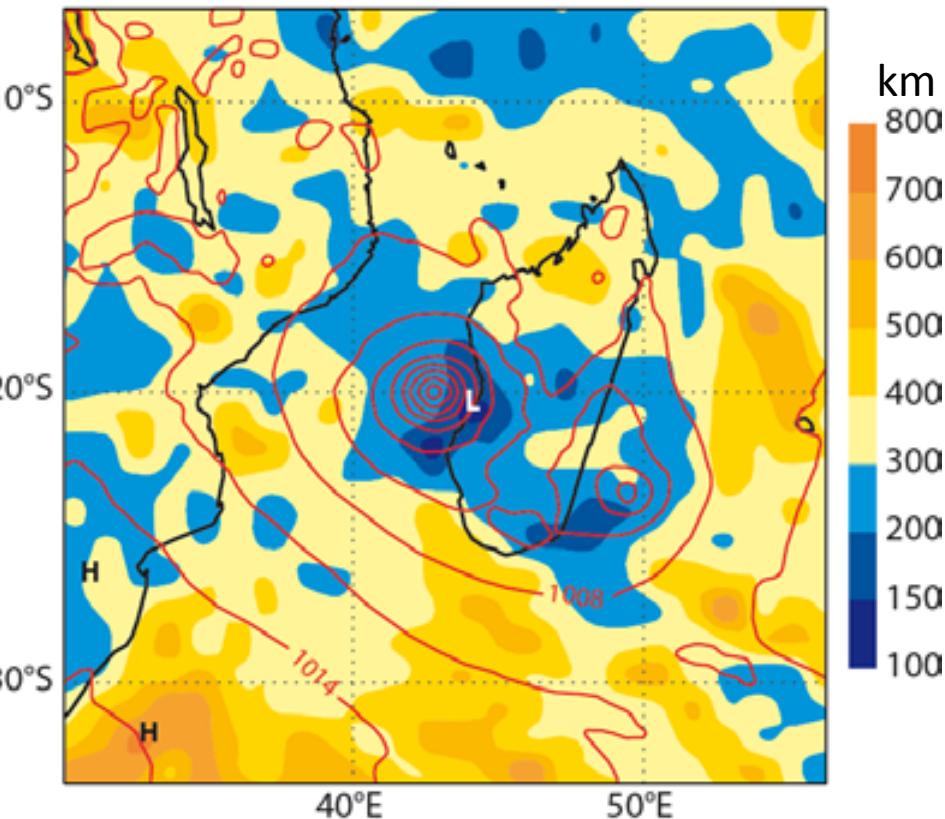


Nov 2013: EDA flow dependent co-variances

Nov 2013: the **EDA size increased from 10 to 25 members**, to provide 4DV-HRES also with flow dependent background error co-variances.
EDA-based perturbations from the past 12 days will be used (sample size=600).



Background error correlation length scale for $\text{long}(p_{\text{msl}})$ and p_{msl}



(From M Bonavita)



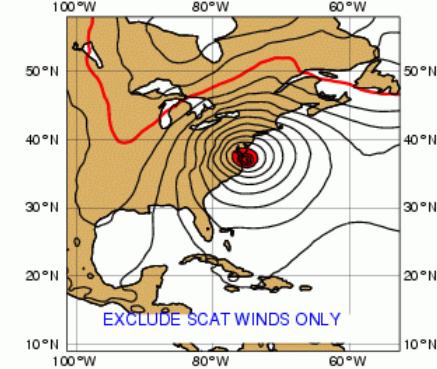
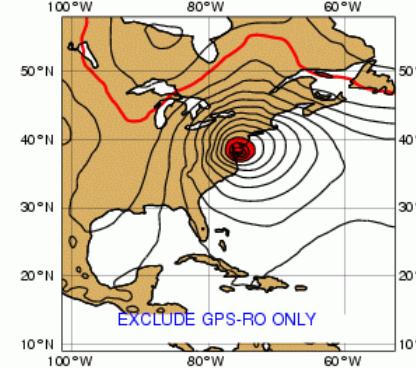
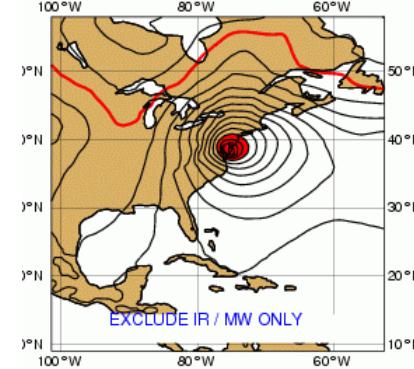
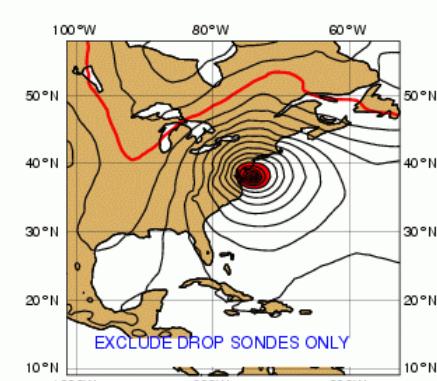
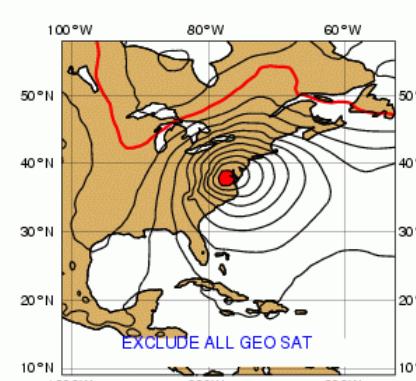
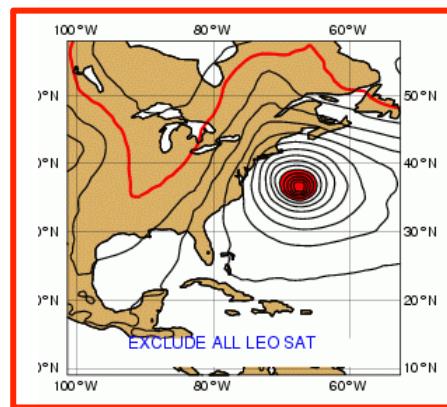
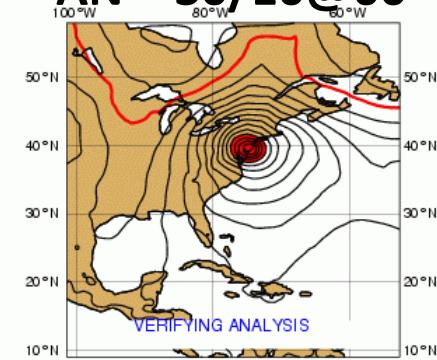
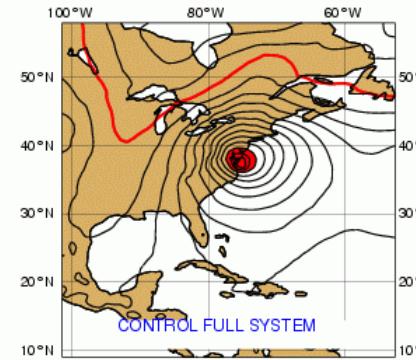
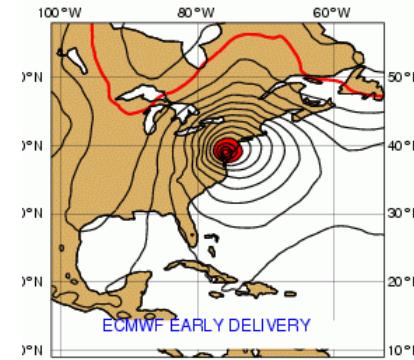
Sandy, OSE and EDA-based flow dependent J_b

AN - 30/10@00

All panels apart
for top-right
shows t+5d fc
from OSE
experiments.

Results indicate
a strong
sensitivity to
Low Earth
Orbiting satellite
data (in red box).

Removing LEO
causes large
errors in +5d fc.

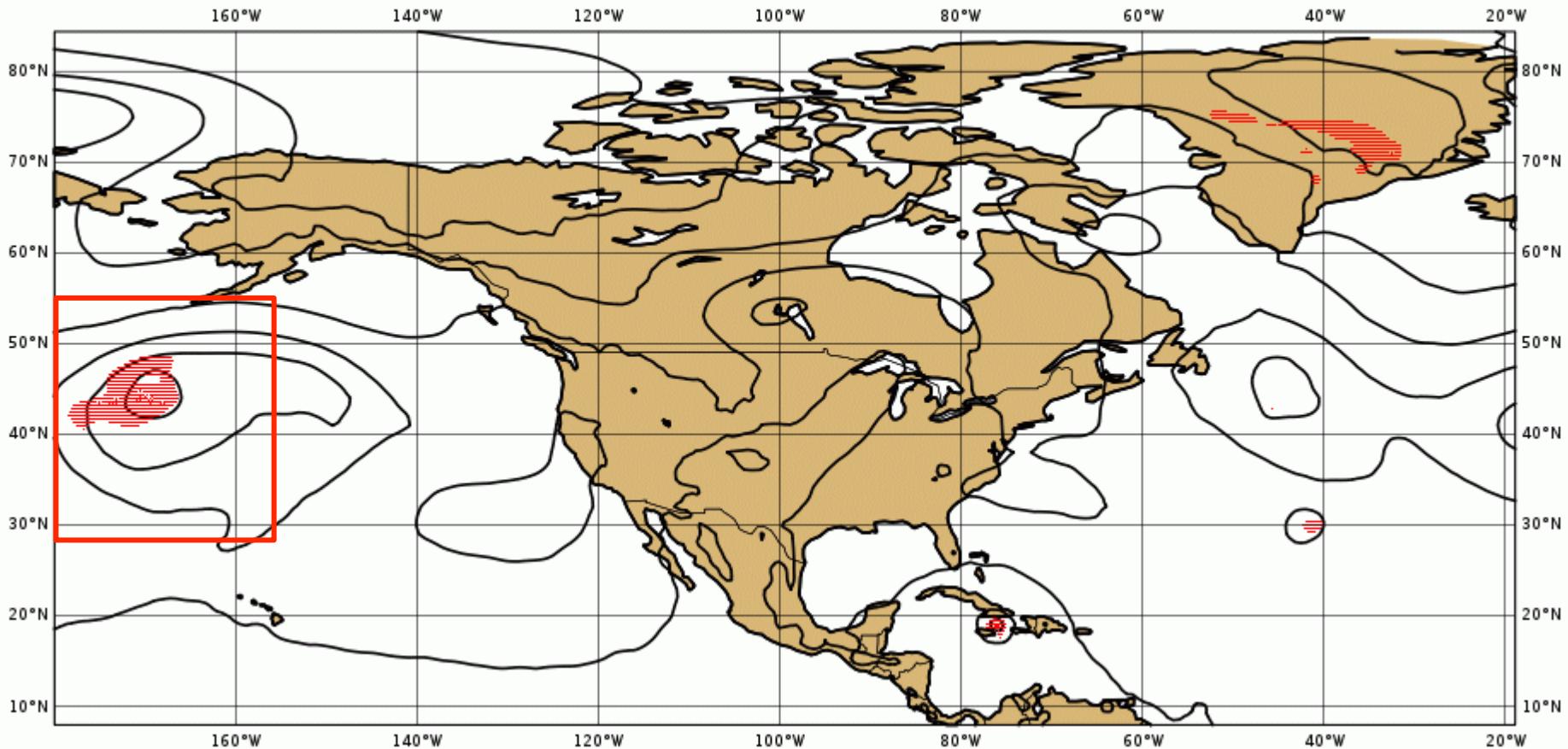


(From T McNally)



Impact of removing LEO data on analysis

Control minus NO-LEO SAT MSLP 2012102500z (after 5 days of data denial)

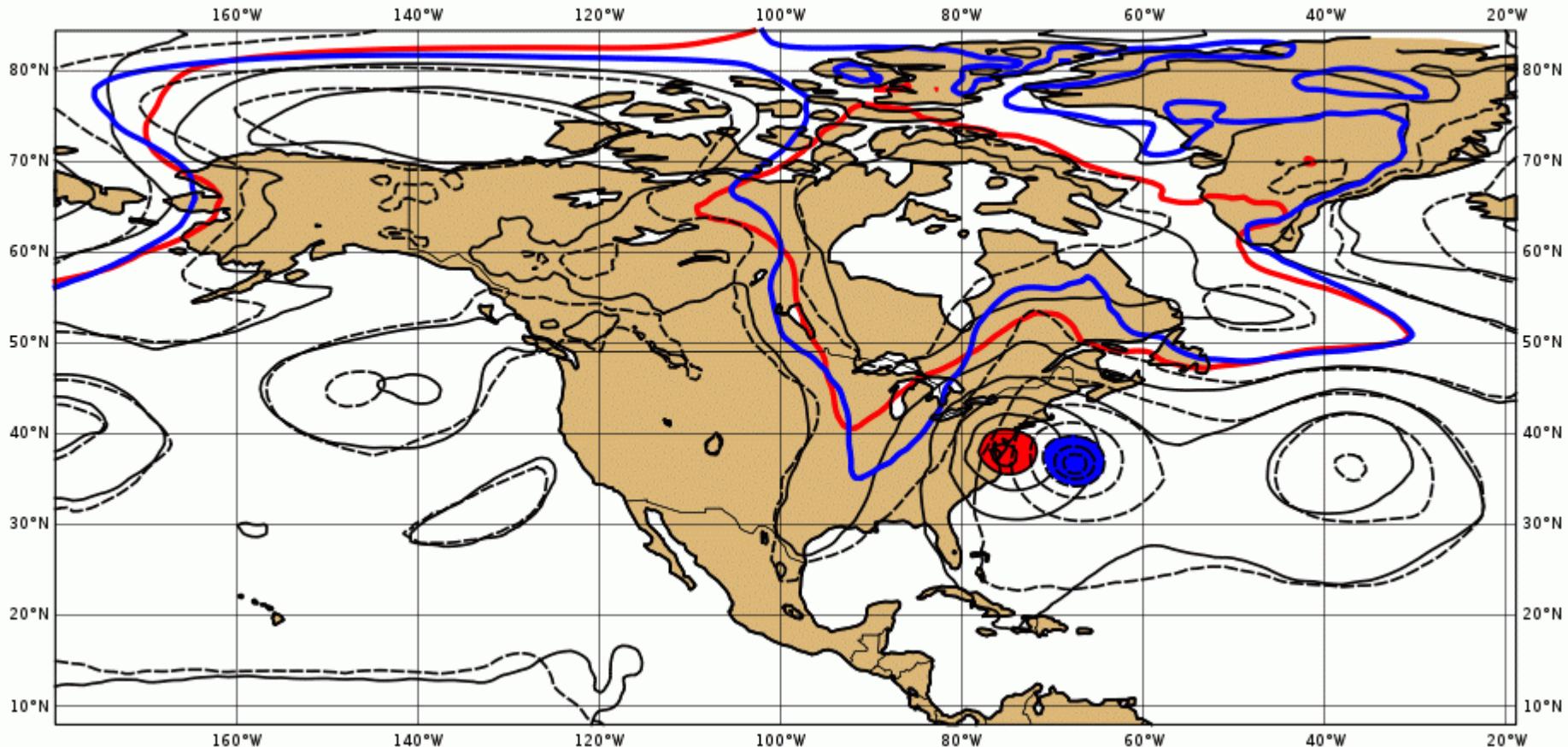


(From T McNally)



Impact of removing LEO data on +5d forecast

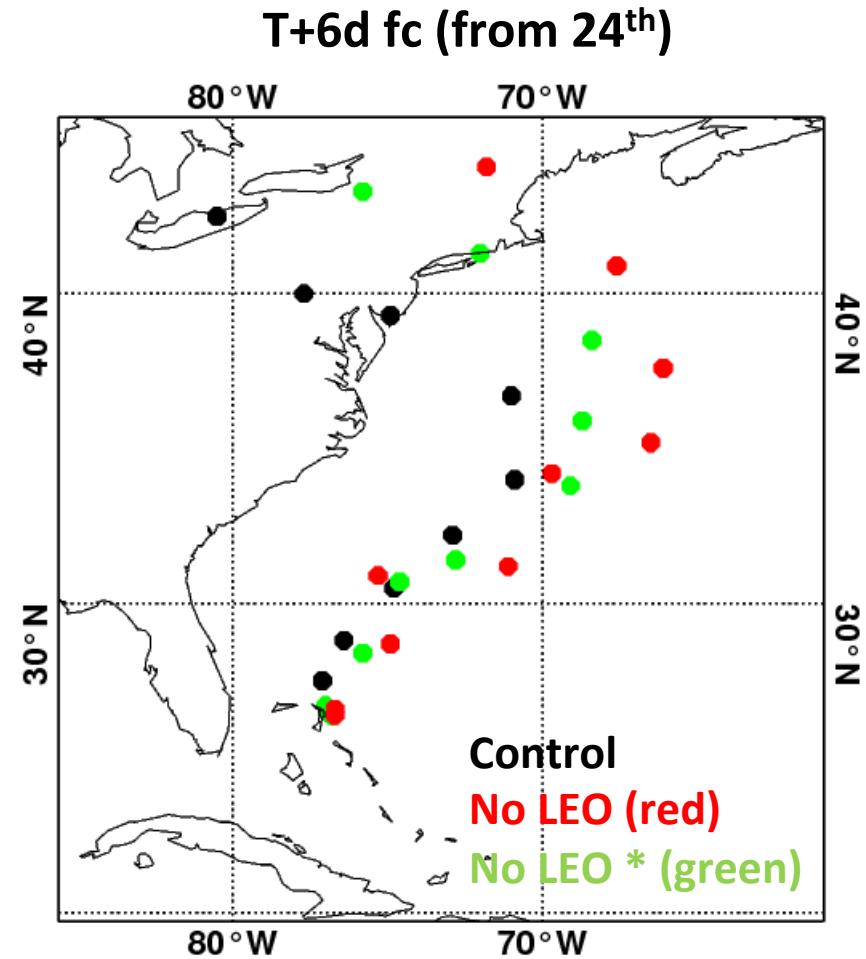
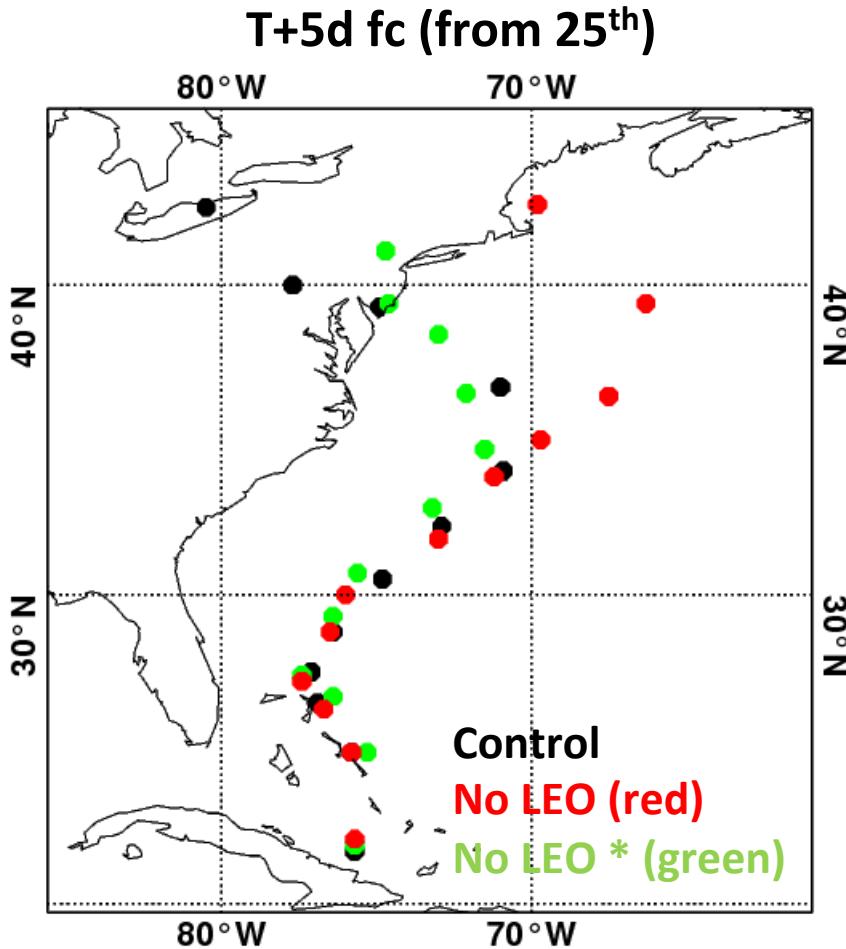
MSLP in Control (red and black solid)
NO-LEO SAT (blue and black dash) VT:2012103000z



(From T McNally)



Impact of using inflated EDA on +5d forecast

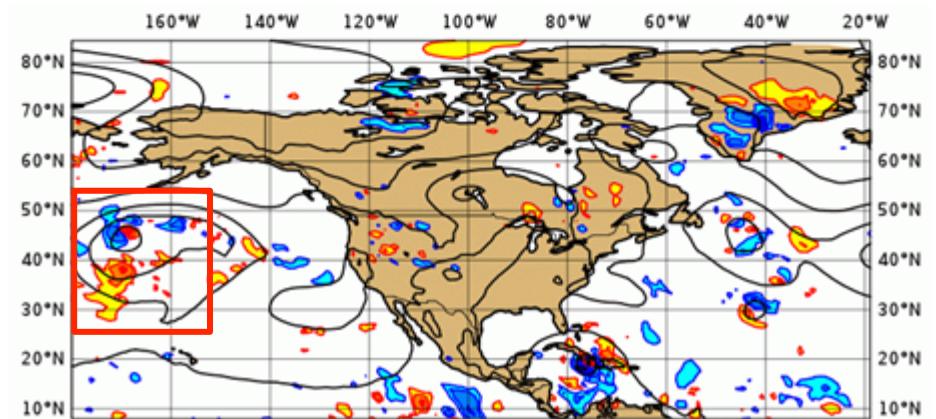
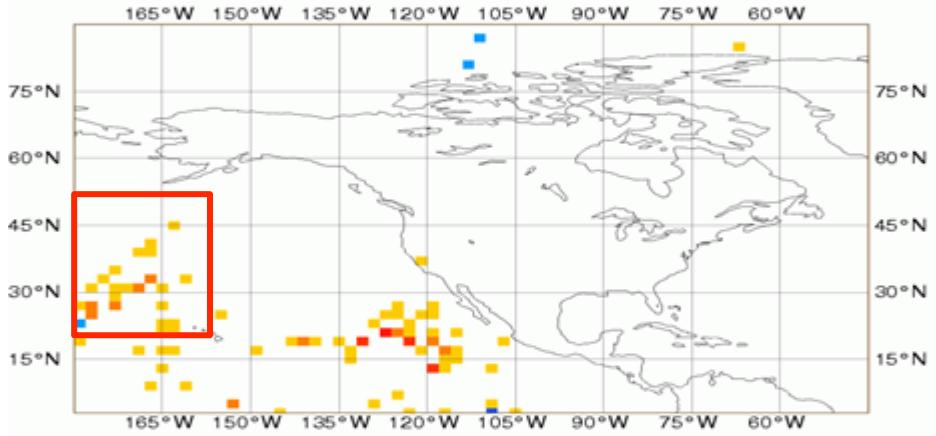
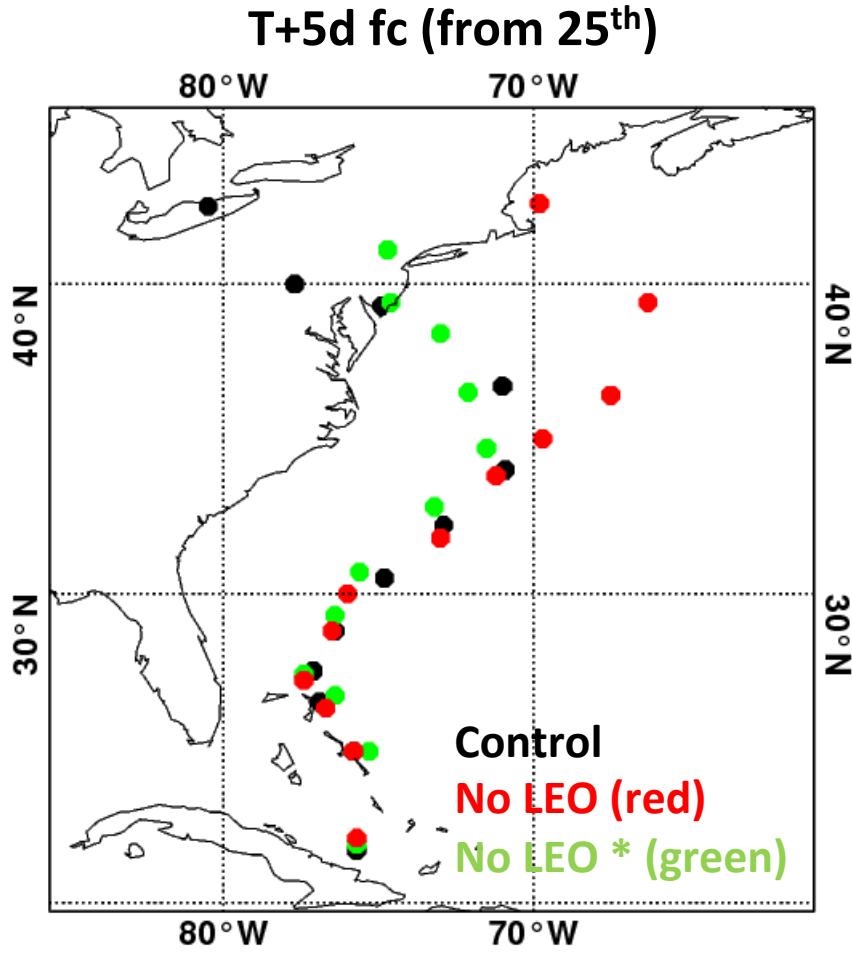


(From T McNally and M Bonavita)



Impact of using inflated EDA on +5d forecast

The track fc improvements with EDA inflation is due to the fact that this leads to the use of extra GEO AMVs data in key area.



(From T McNally and M Bonavita)

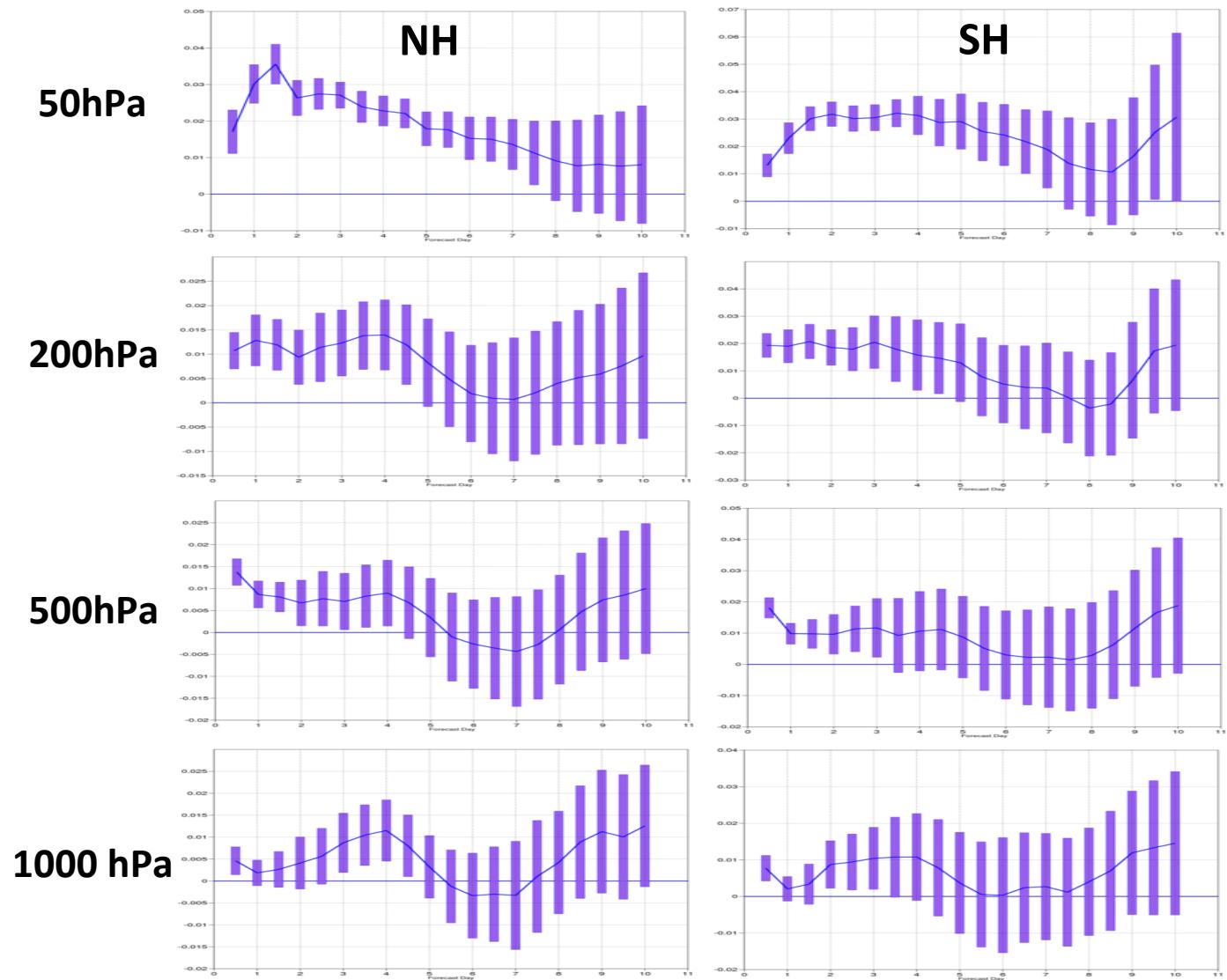


Nov 2013: EDA flow dependent co-variances

Impact of changes implemented in Nov 2013 on the hybrid EDA-4DV.

Plots show of the % of rmse reduction for single forecasts:

- Feb - June 2012
- Resol. T511L91
- 3 Outer Loops (T159/T255/T25)
- Verified against operational analysis



(From M Bonavita)



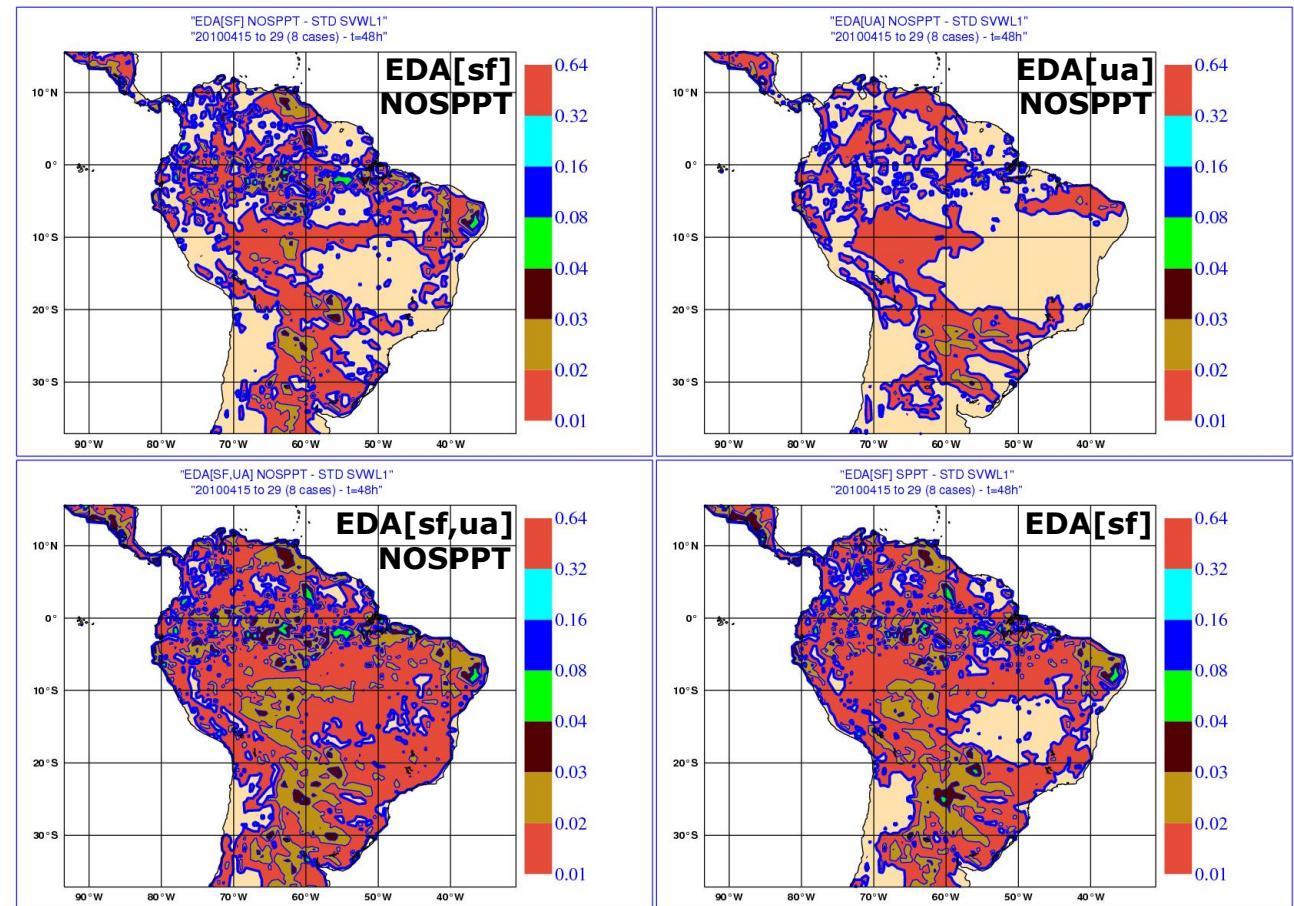
Nov 2013: EDA-based land-surface pert. in ENS ICs

Initial perturbations improved by coupling more the ensemble of analyses and forecasts.

Nov 2013: EDA-based surface initial perturbations were introduced in ENS.

Preliminary results indicate increased reliability in the short forecast range, due to small spread improvements.

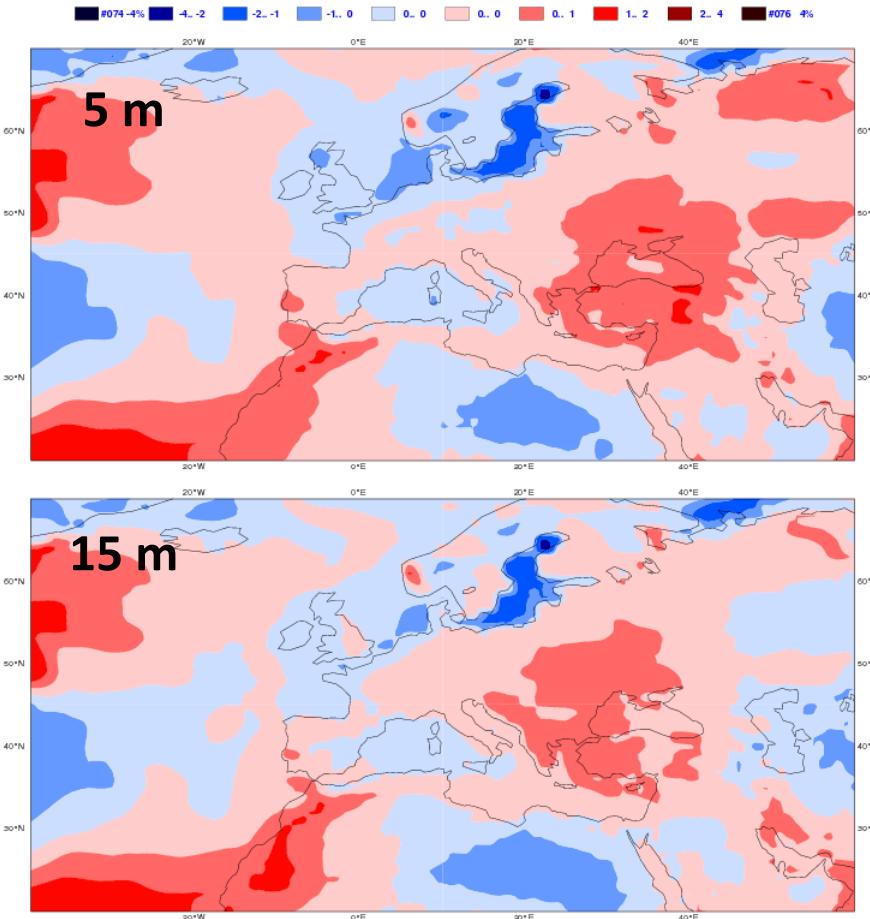
SWVL1 – std($t=48h$)



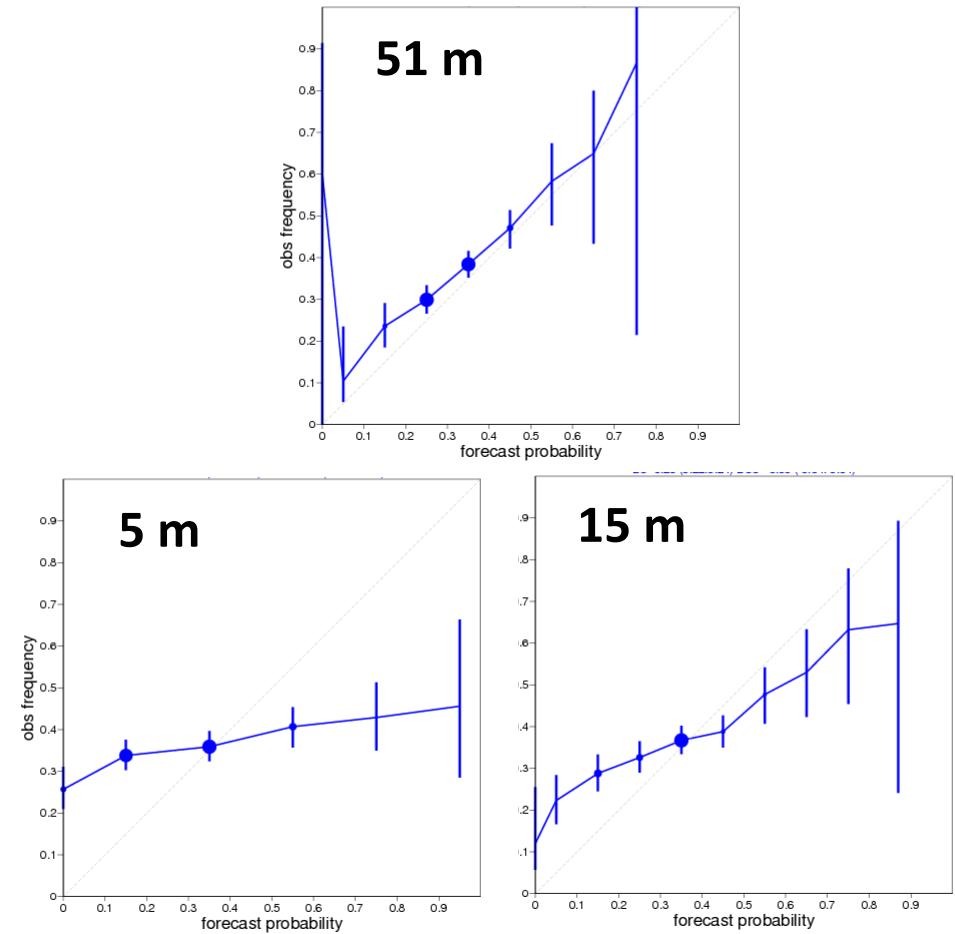


June 2014: extension of refc ensemble (5 to 22 per week)

Impact on calibration 2mtm anomalies – Day 26-32



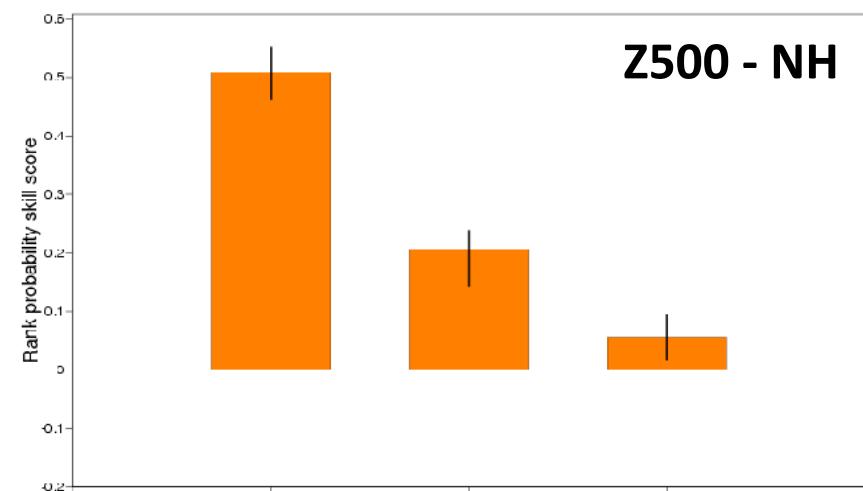
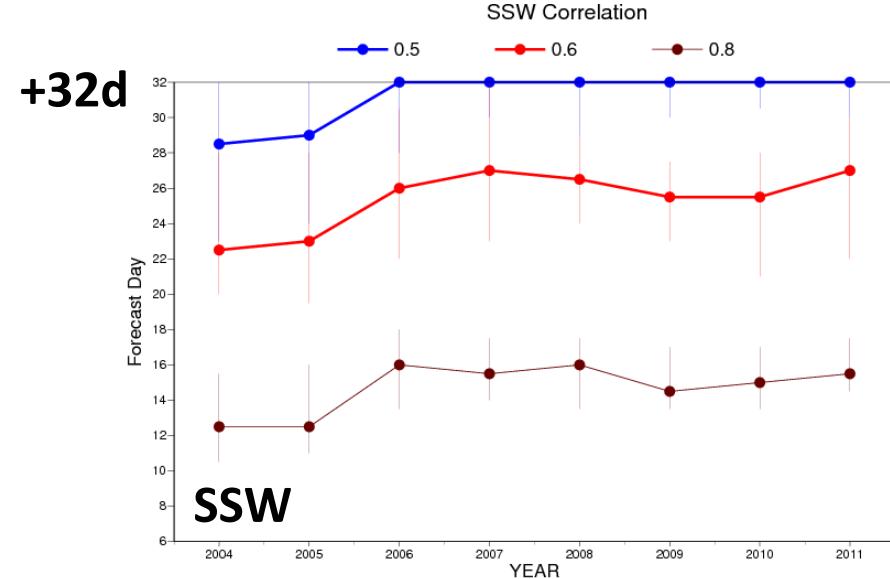
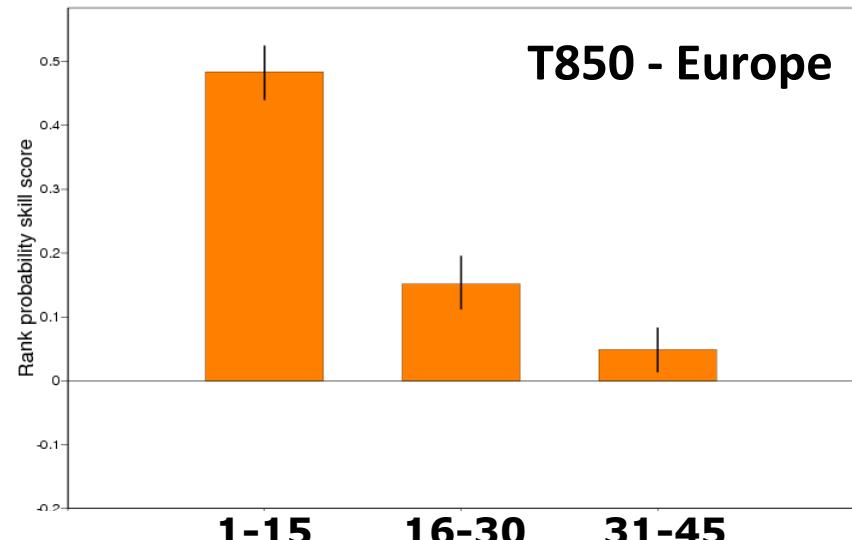
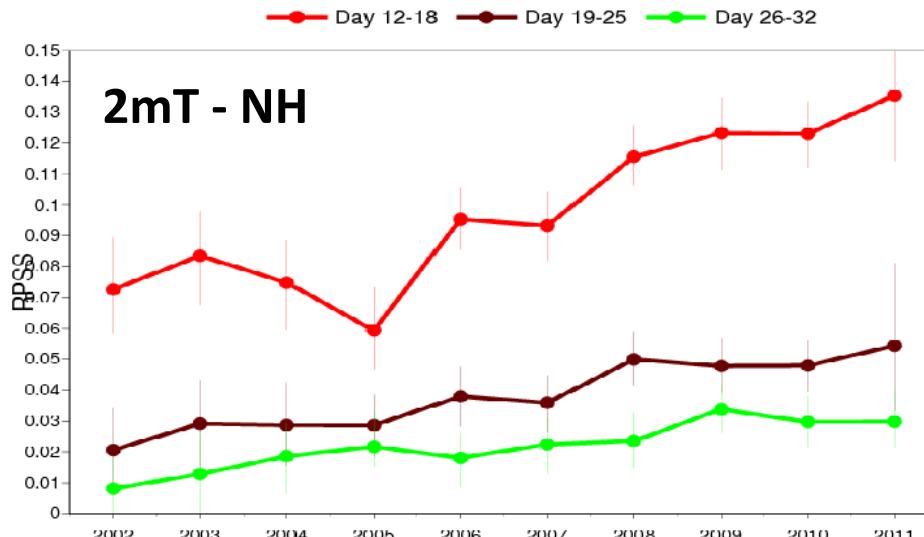
Impact on verification T850- Upper terciles – Week 4





2014/15: ENS extension to 45d/60d to exploit increased skill

2-meter temperature anomalies over the Northern Hemisphere



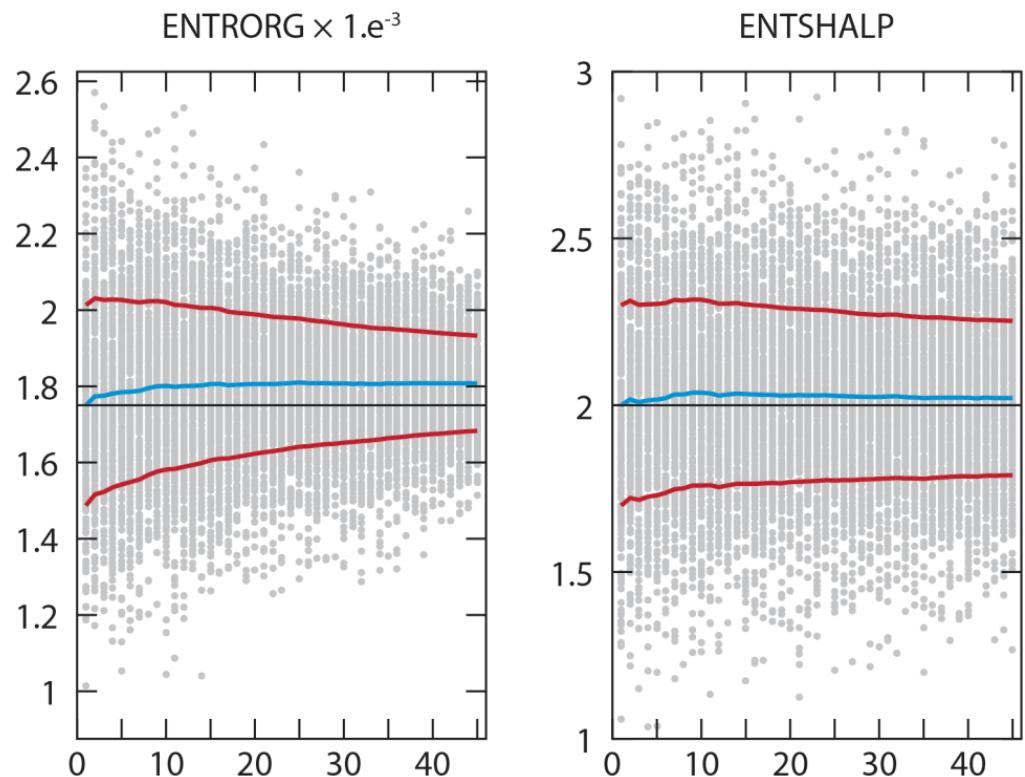


2015/16: more realistic model uncertainty schemes

Model developments should also include improvements to the simulation of model uncertainties.

The plan is :

- i. *To diagnose more thoroughly the operational schemes (SPPT, SKEB);*
- ii. *To assess the potential benefit of making them more realistic;*
- iii. *To explore new ideas in collaboration with other groups (e.g. automatic estimation of model parameters, from FMI)*



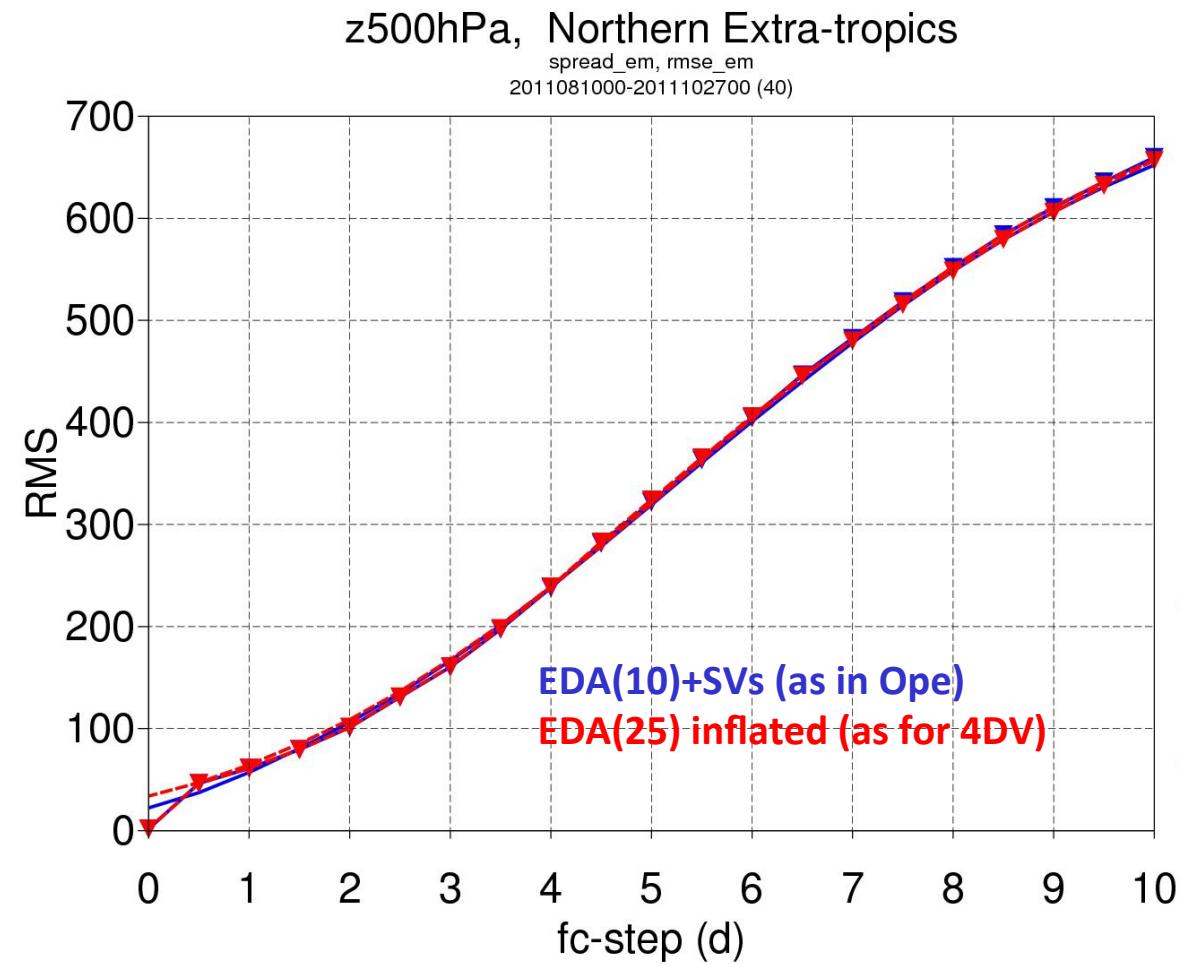
(From P Ollinaho, H Järvinen, P Bechtold, M Leutbecher)



Future: exploit more the EDA to improve ICs

The plan is to assess:

- i. *The impact of increasing EDA size to 25-51 and of increasing the resolution of the EDA outer and/or inner loops;*
- ii. *The possibility to inflate EDA-based perturbation (as done in HRES-4DV) and reduce SV role.*
- iii. *Consider alternative initial perturbations (e.g. from EnKF).*

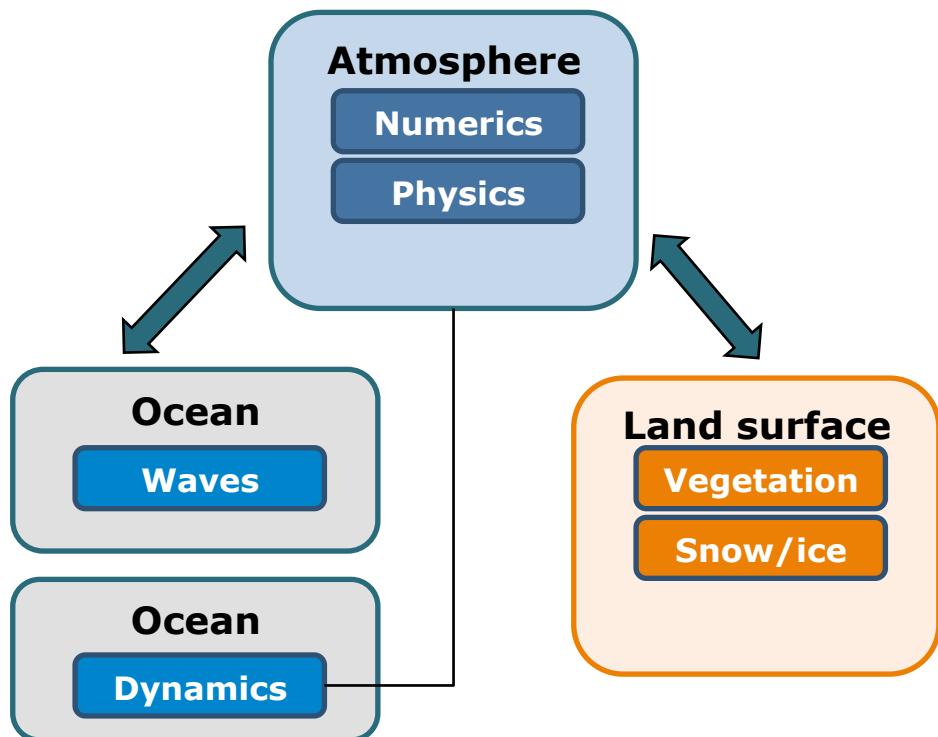


(From S Lang)

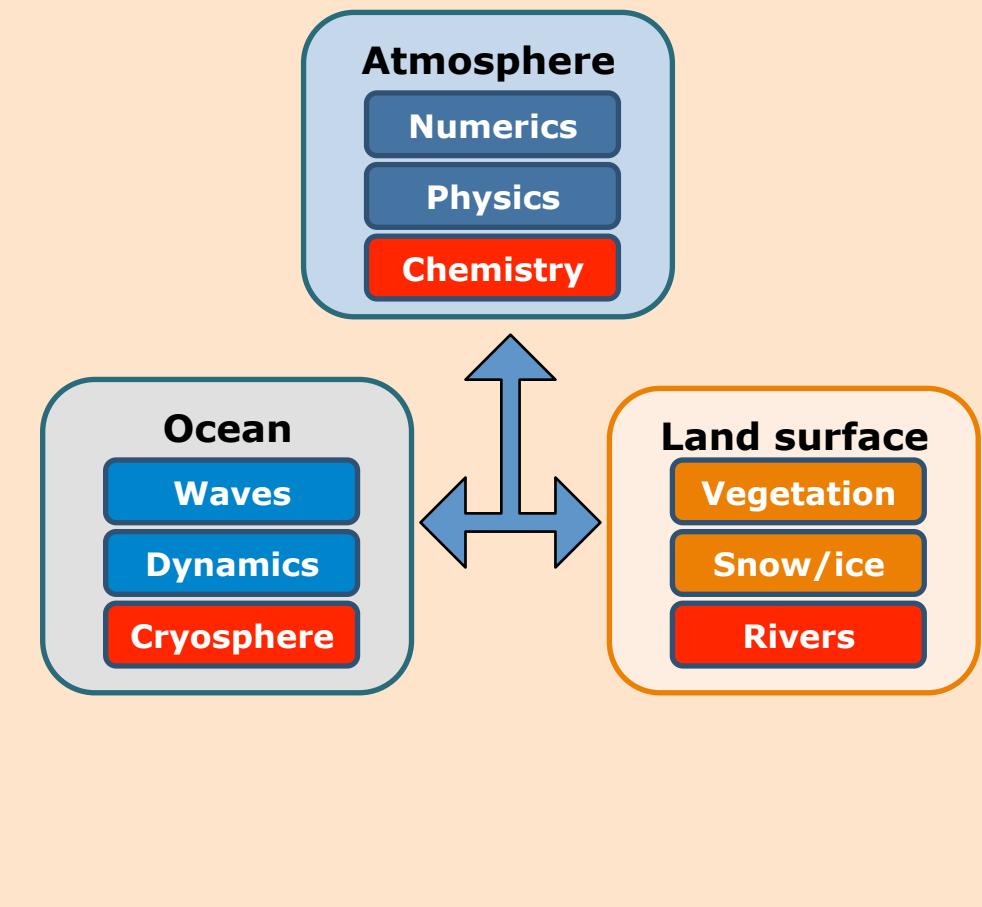


Future: build a more complete model

IFS today



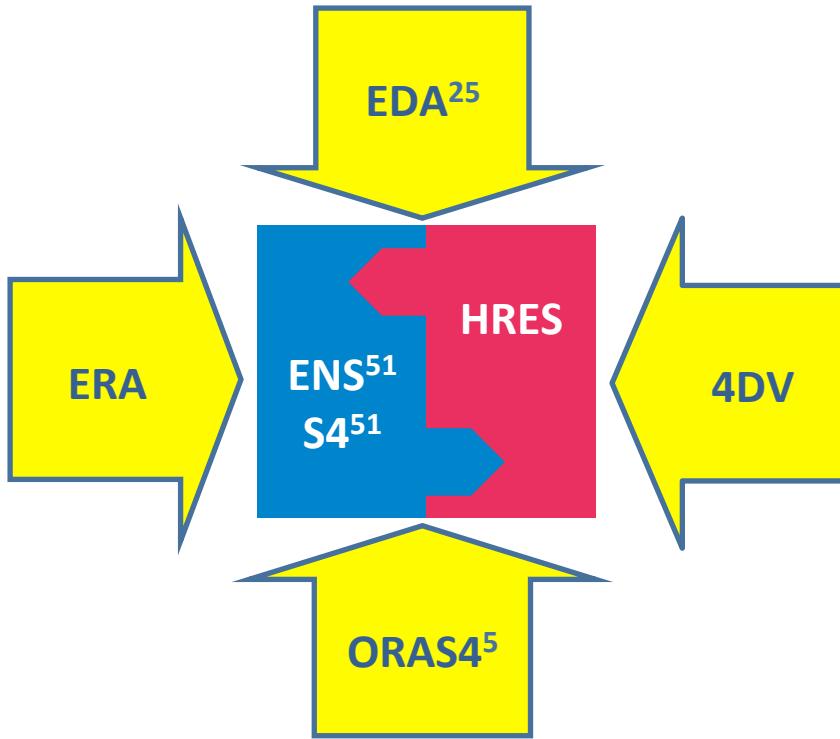
The future





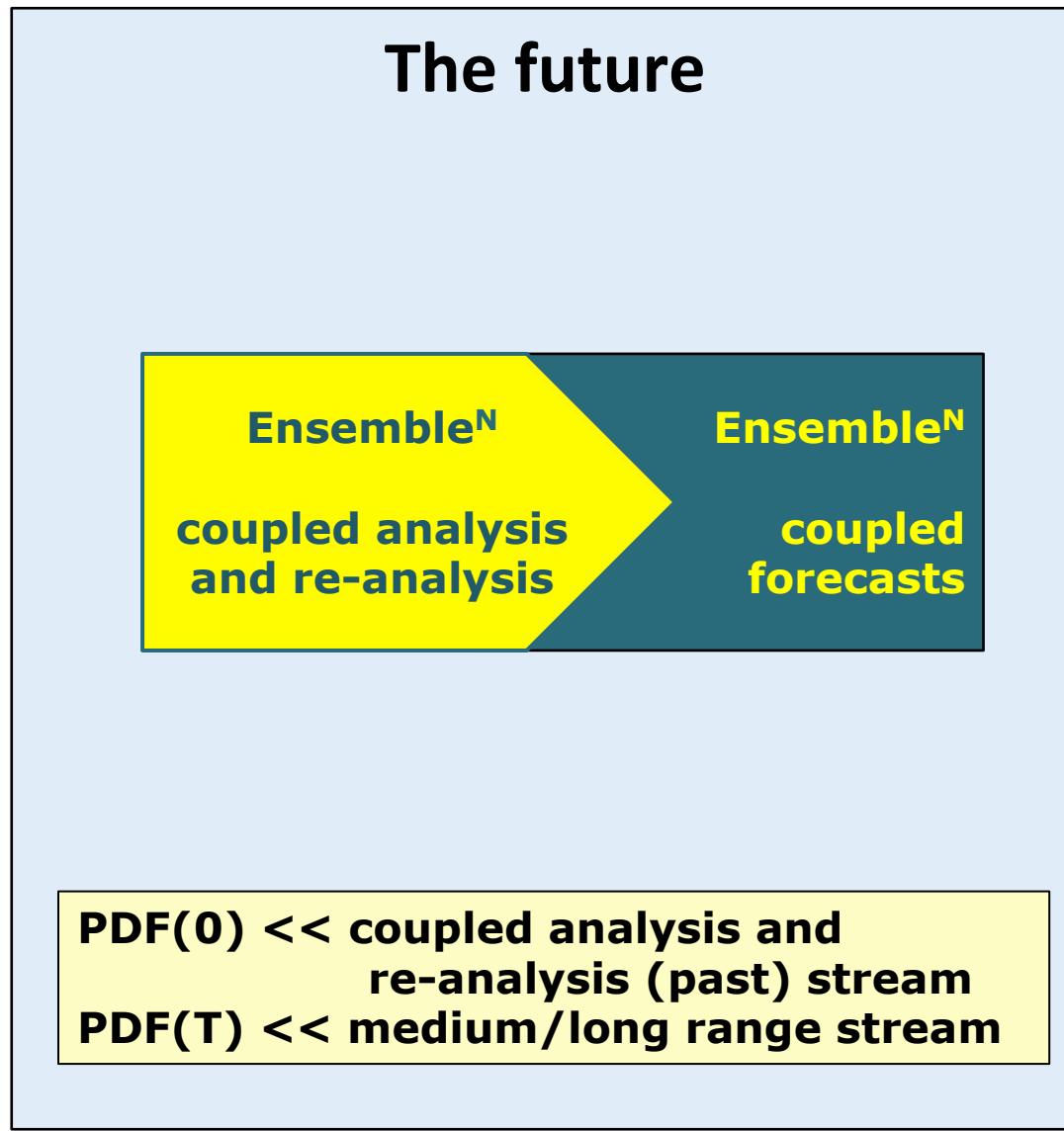
Future: move towards a more integrated approach

Today



PDF(0) << 4DV+EDA+ORAS4
PDF(0) << ERA+ORAS4 (past)
PDF(T) << HRES+ENS/S4

The future



**PDF(0) << coupled analysis and
re-analysis (past) stream**
PDF(T) << medium/long range stream



Conclusions

- 20 years ago ECMWF was issuing one forecast only up to 10 days.
- Today, ensemble methods are used in analysis and forecast mode, and ECMWF issues probabilistic forecasts up to 13-month, and ensemble-based analysis error estimates.
- In 2014/15 we will upgrade the medium-range/monthly ensemble: improve the marine component (coupling waves <> currents <> sea-ice), increase the resolution (to ~ 25 km in atmosphere and ocean)
- In 2016/17 we will upgrade the seasonal ensemble
- The future will see ECMWF providing **more accurate estimates of the most likely scenario and its uncertainty at analysis and forecast time for all variables** (atmospheric, land, ocean, chemical and composition, ...).



... extra slides



How do ensembles perform? Scores' trends CRPSS TP24 EU

Similar, although smaller in size, improvements can be seen by looking at 24h total precipitation, with skill gains of about 1.75 days between 2001 and 2012.

The plot shows the forecast lead time when CRPSS crosses a 10% value.

ECMWF EPS 12UTC forecast skill

total precipitation

Continuous ranked probability skill score

Europe (lat 35.0 to 75.0, lon -12.5 to 42.5)

— 12mMA of CRPSS reaches 0.1

